Essays in Applied Microeconomics and Behavioral Finance

Marco Giovanni Nieddu

TESI DOCTORAL UPF / ANY 2018

DIRECTOR DE LA TESI
Albrecht Glitz - Departament de Economia i Empresa

To my parents

Acknowledgements

I am deeply grateful to my supervisor, Albrecht Glitz, for his guidance, patience and suggestions throughout my entire PhD. For the works presented in this thesis, and for what I have learned by working with them, I also have a large debt of gratitude to Gianmarco León and Matteo Bobba who showed me how to do research.

I need to thank the whole Department of Economics and Business of the Universitat Pompeu Fabra, with a special acknowledgement for Jose Apesteguia for his invaluable help. I also need to say a sincere thank you to Marta Araque and Laura Agustí. Their patience and efforts really made a difference to me.

I thank my favorite colleagues and friends Ana, Cristina, Darren, Francisco, Maria, Marta, Miguel, Patrick, and Tomas, to whom I wish the brilliant career and happy life they deserve. Also, I owe a special dedication to all the faculty members of the Paris School of Economics, and to the good friends I made during my year in Paris: Christina, Laura, Leo, Kenneth, Paola and Rozenn. There are honestly no words to describe how lucky I am to have met Lorenzo Pandolfi – who also became a fantastic coauthor – and Niklas Heusch.

Last, but foremost, I am deeply indebted to my parents who instilled in me the love for beauty and knowledge.

Hic Abundant Leones

Abstract

This thesis consists of three chapters. In the first chapter, I investigate how promotion incentives affect the performance of high-skilled public employees. I study a centralized evaluation process awarding the eligibility for associate and full professorship in Italian academia, and show that the perspective of a promotion induces scholars to increase their research productivity. In the second chapter, I present the results from a laboratory experiment designed to assess whether and how financial literacy influences the way individuals perceive and evaluate financial assets. By comparing participants' investment decisions under different treatments, I show that the lack of financial literacy lowers the subjective value that investors assign to risky financial assets. The third and last chapter is devoted to an empirical analysis of the link between university quality and employment opportunities. I find that postgraduate students who receive incentives to attend higher-ranked universities are more likely to be employed one year and a half after the end of their studies.

Resumen

Esta tesis consiste de tres capítulos. En el primer capítulo, analizo cómo los incentivos de promoción afectan el rendimiento de empleados públicos altamente calificados. Estudio un proceso de evaluación centralizado que determina la elegibilidad para posiciones de profesor titular y catedrático en la academia italiana, y encuentro que aquellos académicos que tienen la posibilidad de obtener una promoción aumentan la productividad de su investigación. En el segundo capítulo, presento los resultados de un experimento de laboratorio diseñado para determinar si y cómo la alfabetización financiera afecta la manera en la cual los individuos perciben y evalúan los productos financieros. Al comparar las decisiones de inversión de participantes sujetos a distintos tratamientos, muestro que la falta de alfabetización financiera afecta el valor subjetivo que los inversores le atribuyen a activos financieros riesgosos. Finalmente, el tercer capítulo analiza empíricamente la relación entre la calidad universitaria y las oportunidades de empleo. Encuentro que los estudiantes de posgrado que reciben incentivos para asistir a universidades de alto nivel tienen mayor probabilidad de estar empleados un año y medio después de concluir sus estudios.

Preface

The present thesis consists of three essays. The work presented in the first chapter – *The Effectiveness of Promotion Incentives for Public Employees: Evidence from Italian Academia*, coauthored with Lorenzo Pandolfi – investigates how promotion incentives affect the productivity of high-skilled public employees. In a fuzzy regression discontinuity design we exploit the three bibliometric thresholds of the 2012 Italian National Scientific Qualification, the centralized evaluation procedure awarding the eligibility for career advancements in public universities. Specifically, we compare the 2013-2016 research performance of assistant professors who barely qualify for associate professors with that of candidates who barely miss the qualification. While the former are incentivized to enrich their publication record in order to meet the expected future thresholds for the full professor qualification, the latter must re-apply for the associate professor qualification, thus facing lower promotion thresholds. We find that barely qualified scholars publish significantly more papers – and in journals with comparable quality – than their unsuccessful colleagues. The relationship between the increase in publications and the distance from the expected threshold for full professorship is inverted-U shaped: promotion incentives are mostly effective when the promotion threshold is neither too difficult nor too easy to be met. Our results emphasize the importance of promotion incentives as an effective tool for public management to enhance the productivity of state personnel. Additionally, given our focus on academia, they shed novel light on the responsiveness of scholars to publication-based hiring and promotion schemes. This issue is particularly important as the production of knowledge is a key driver of economic growth.

In the second chapter – *Cutting Through the Fog: Financial Literacy and the Subjective Value of Financial Assets*, coauthored with Lorenzo Pandolfi – we examine the impact of financial literacy on investors' subjective valuation of financial assets. In a laboratory experiment with a two-by-two design, we study how the certainty equivalent of a risky lottery changes when varying the framing of the lottery and participants' level of financial literacy. We find that framing the lottery as a financial asset rather than as a simple coin toss reduces the average certainty equivalent by approximately 20% and lowers participants' understanding of the lottery's structure. Enhancing financial literacy by teaching basic financial notions offsets the negative effects of the financial framing, as it improves respondents' understanding of the lottery and increases the average certainty equivalent. Our results – which can be rationalized by ambiguity aversion – shed new light on the linkages between financial literacy and financial decision-making. Additionally, they highlight the importance of promoting financial education to stimulate households' financial market participation.

Lastly, in the third chapter – *Students' Choices and The Employment Returns*

of University Quality – I study how incentives rewarding investments in education quality affect the university choices of a sample of Italian postgraduate students as as well as their later employment opportunities. I exploit a budget tightening that occurred between two editions of a regional program providing financial support to students who enroll in master's degree programs. The reduction in the number of scholarships available translates into an exogenous increase in competition. Since applications are ranked on the basis of both applicants' curriculum vitae and the score of the chosen university in a popular university ranking, the treated cohort is incentivized to choose higher-ranked institutions. Moreover, students lacking cv points receive the strongest incentive. I compare – in a difference-in-differences framework – the university choice and the post-master employment outcomes of applicants belonging to different cv groups before and after the increase in the competition for scholarships. I find that the treated cohort of applicants chooses universities which have a score 8 points higher (out of 60). The increase in university quality is stronger for the group of students who are neither too far from the (expected) thresholds for winning the scholarship, nor too close. After completing their studies, this group is around 30 percentage points more likely to be employed 15 months after the end of the master's degree. As incentives induce a switch from universities in Central and Southern Italy to universities in the North – where the unemployment rate is typically lower – I suggest that part of the employment effect might be driven by students moving to areas characterized by a more favorable labor market.

Index

List of Figures

List of Tables

Chapter 1

THE EFFECTIVENESS OF PROMOTION INCENTIVES FOR PUBLIC EMPLOYEES: EVIDENCE FROM ITALIAN ACADEMIA

1.1 Introduction

Rewards for good performance are a key tool for firms and organizations to motivate and retain employees. Numerous studies find performance-based incentives to be largely effective in the private sector.[1] Conversely, the literature on how to motivate civil servants is much smaller and limited mostly to pay-for-performance schemes, with very few studies focusing on promotion incentives. This lack of empirical evidence is concerning, as career-based incentives, in fact, represent the main motivational lever in the hands of public management.[2] Compensation schemes in the public sector are indeed typically rigid and do not easily allow for

[1]The literature on the effectiveness of financial incentives in the private sector includes, among others, Lazear (2000), Gaynor et al. (2004), Shearer (2004) and Friebel et al. (2017). For career-based incentives, see Kwon (2006) and Campbell (2008). See also Lazear and Oyer (2013) for an exhaustive review.

[2]Additionally, according to Haeck and Verboven (2012), promotions constitute the largest source of incentives for public workers. Studying the case of an European university, they find that its labor market is characterized by a strong barrier at the entry level, salaries that evolve independently from external wages and long internal career progressions.

the inclusion of discretionary performance-based components (Finan et al., 2017). Additionally, in many developed countries, a significant share of the entire labor force is employed in the public sector, whose productivity thus represents a key determinant of economic growth.[3] Deeper insight into whether and how promotions can effectively incentivize workers in public organizations is thus needed, as the question is important from both a policy perspective and an academic perspective.

The aim of this paper is precisely to shed light on this topic. We assess the impact of promotion incentives on the performance of high-skilled public employees by studying whether a quasi-random assignment of different career prospects affects the productivity of approximately 5,000 assistant professors in Italy. Specifically, we exploit the introduction, in 2012, of a centralized evaluation procedure awarding the eligibility for career advancements – namely, the *National Scientific Qualification* (henceforth NSQ) – on the basis of past performance. We take advantage of a peculiar feature of this procedure: success in the NSQ depends on scholars' past research productivity, measured by three *bibliometric* indicators that are required to be above certain observable and well-defined thresholds. Hence, in a regression discontinuity design with three running variables, we compare the post-2012 research productivity of barely successful and unsuccessful assistant professors applying for the associate professor qualification. While the former can achieve the qualification for a full professorship in the subsequent round of the NSQ, the latter first need to re-apply for the associate professor qualification.[4] Success or failure in the 2012 NSQ thus generates very different promotion incentives. Qualified candidates have the incentive to enrich their publication records in order to meet the higher eligibility requirements for a full professorship in the subsequent round of the NSQ. Conversely, the goal for barely unsuccessful scholars remains meeting the associate professor thresholds, which are, by definition, very close.[5]

Our (triple) regression discontinuity estimates show that achieving the qualification in 2012 – and thus being exposed to higher promotion incentives – has a

[3] As of 2015, public employment accounts for 18.1% of total employment across OECD countries. In Scandinavian countries, this share increases up to approximately 30%, almost twice that of the US (15.3%) (OECD, 2017)

[4] Although the system does not explicitly prevent assistant professors from applying directly for the full professor qualification, the probability of succeeding without having already obtained the associate professor qualification is *de facto* very unlikely, as we show in greater detail in Section 1.4.

[5] By 'unsuccessful scholars', we refer to both the applicants who are denied the qualification and those who withdraw their application before the committee evaluates their applications. In a regression discontinuity framework, the 'barely unsuccessful' candidates are the ones whose *bibliometric* indicators are almost at the threshold. Thus, the gap they must fill in order to overcome the minimum requirements in one of the later rounds of the NSQ is tiny.

positive and significant effect on the number of scientific papers published in the subsequent four years. The marginally qualified scholar publishes on average 6 items more than her marginally non-qualified colleague. This effect is sizable, as it corresponds to a 38% increase with respect to the average number of publications in the entire sample. We also find that the average publication quality – proxied by different measures of the journal's prestige – does not exhibit any discontinuity at the multidimensional threshold. The increase in the quantity of publications thus does not occur at the expense of their average quality. Additionally, we provide suggestive evidence that qualified scholars tend to expand their co-author network and receive more citations compared to their unsuccessful colleagues. When investigating the heterogeneity of our results depending on candidates' gender, we find no evidence of a differential responsiveness to the provision of promotion incentives between male and female assistant professors. However, we find that women who comply with the cutoff rule are less likely to achieve the qualification than their male colleagues, which may suggest the presence of gender discrimination in the evaluation procedure.

Digging deeper into the distributional effects of promotion incentives, we show that the effect of the qualification on the number of publications is heterogeneous, depending on scholars' distance from the (expected) future promotion thresholds. The relation between the increase in productivity and the distance from the thresholds for the full professor qualification is indeed inverted-U shaped: the effect is the strongest for candidates in the middle of the distribution, while it is not statistically significant in the group of candidates for whom the minimum requirements to earn the qualification for a full professorship are either too close or too far. This result is consistent with the view that incentives are mostly effective when "the promotion is possible, but neither too hard to achieve, nor too easy" (Lazear and Gibbs, 2014, p.269). Besides shedding light on the distributional consequences of the promotion incentives, this evidence lends important support to our empirical strategy. Most of the competing explanations for our main results – for instance, a motivational effect arising from succeeding in the NSQ, or other changes in scholars' daily life occurring after the achievement of the qualification – can hardly be reconciled with the heterogeneity of the estimated effects depending of the distance from the future promotion threshold.

Several additional pieces of evidence corroborate the robustness of our results and confirm that promotion incentives are the main driver of our findings. First, we take advantage of the longitudinal dimension of our data to rule out that the observed discontinuity in the post-call research productivity is driven by a decline in performance of discouraged, unsuccessful candidates rather than by an increase in the productivity of qualified scholars. Second, we replicate our analysis with the sample of associate professors applying for the full professor qualification in 2012. Consistent with the fact that the promotion incentives vanish once the top

ladder of the academic hierarchy is reached, we do not observe any discontinuity in this alternative sample. Further, we show that the increase in productivity of qualified candidates occurred already in 2013, immediately after the achievement of the NSQ and prior to the actual promotion to associate professor. Hence, the timing of the effect suggests that our results are not driven by a variation in teaching duties, research funds, and other aspects of scholars' routine that may change with career advancement. Lastly, we exploit the between-field heterogeneity in the share of candidates qualified for an associate professorship on the total number of associate professors already on staff to show that competition for vacancies is not driving our findings.

This study contributes mainly to the personnel economics literature and, more precisely, to the stream of studies focusing on the design of incentives in the public sector. As highlighted by Finan et al. (2017), public sector pay schemes are typically flat, with salaries that are mechanically determined by seniority and position and rarely linked to workers' performance. Precisely because of this, most studies focus on programs implemented in the context of randomized control trials rather than on actual government policies and typically do not involve high-skilled workers. For instance, Muralidharan and Sundararaman (2011) evaluate the impact of a randomized performance-pay program in India and find that linking teachers' pay to students' test scores has a positive effect on learning. In another randomized experiment, Duflo et al. (2012) show that performance-related pay lowers absenteeism among Indian teachers, which in turn translates into better students' performance.[6] This literature also highlights a potential pitfall of implementing performance-pay schemes in the public sector: as public jobs typically involve multiple tasks, financial incentives based on the performance in a specific task can reduce workers' effort in another (Baicker and Jacobson, 2007; Glewwe et al., 2010).

Therefore, while some studies have examined the effectiveness of monetary incentives on the productivity of public employees, the existing literature has overlooked the role of promotion incentives. The latter constitute a complicated subject of study – even in the private sector – since promotion incentives can hardly be implemented in the context of randomized controlled trials, and quasi-experimental evidence on the topic is rare. To our knowledge, the only analysis of the relationship between promotion incentives and workers' productivity in the public sector is that by Karachiwalla and Park (2017).[7] The authors exploit the Chinese system regulating teachers' career advancement to test the prediction of

[6]Other studies focusing on the role of financial incentives in the public sector are those by Lavy (2002), Gertler and Vermeersch (2013), Dal Bó et al. (2013) and Olken et al. (2014). See Finan et al. (2017) for an exhaustive review.

[7]The literature on promotion incentives in the private sector is comparatively more extensive, but also limited. See, for instance, Kwon (2006) and Campbell (2008).

a tournament model of promotions and show that promotion incentives are associated with higher levels of performance.[8] Our study is therefore the first that exploits a quasi-experimental variation in promotion incentives – coming from explicit promotion thresholds implemented within an actual governmental policy – to study how these affect the productivity of a large sample of high-skilled public employees.

More broadly, this paper is also related to the recent literature focusing on the centralized evaluation systems that have been introduced in the last decade in several European countries to regulate access to public university positions. Similar to the Italian NSQ are, among others, the *Acreditación* in Spain and the *Habilitation à diriger des recherches* in France. All the related studies focus on the functioning of the evaluation process and, more specifically, on the role of gender (Bagues et al., 2017; De Paola and Scoppa, 2015; De Paola et al., 2017) or of direct connections between evaluators and candidates (Zinovyeva and Bagues, 2015). None of them examines the potential implications for scholars' productivity, as we do in this study.

Additionally, besides representing an ideal framework for our analysis – with observable measures of individual productivity, a well-defined promotion system, and a clear hierarchical structure – academia is a setting of particular interest *per se*. The production of knowledge is indeed recognized to be one of the main engines of economic growth and, in many countries, the main provider of research and education is the state. However, while there is some evidence on the determinants of productivity in the scientific production process (Waldinger, 2012; Moser et al., 2014; Borjas and Doran, 2015; Waldinger, 2016), little is known about how academics respond to different recruitment and promotion schemes. We show that explicitly linking career advancements to scholars' publication records can effectively foster their scientific production.[9]

The rest of the paper is organized as follows. Section 1.2 describes the regulatory framework and the key features of the NSQ. The data used for the empirical analysis and the identification strategy are reported in Sections 1.3 and 1.4, respectively. We then present the first-stage estimates in Section 1.5. Section 1.6 contains our main results, together with some additional robustness tests. In Section 1.7, we dig deeper into the promotion incentives mechanism by exploring the the distributional effects of promotion incentives and discarding potential competing hypotheses for our findings. Finally, Section 1.8 concludes.

[8] Additionally, Ashraf et al. (2018) study the role of career prospects on the selection of workers in the public sector (health) in Zambia and on the quality of the service delivered. However, this study focuses on the selection channel.

[9] In this sense, our study also speaks to the literature on the effects of achieving tenure on scholars' research productivity. See Faria and McAdam (2014) and references therein.

1.2 Institutional Setting

In 2010, the Italian Ministry of Education, University and Research (MIUR) deeply reorganized the public university system through the so-called Gelmini reform. The latter introduced a new recruitment and promotion system regulating the access to the two top ranks of the academic hierarchy: the associate professorship and the full professorship.[10] Until that time, the hiring and the promotion processes were fully decentralized and each academic department had complete discretion over the selection procedure. Since the reform came into force, however, earning an associate or full professorship is conditional on having achieved a qualification – the National Scientific Qualification (*Abilitazione Scientifica Nazionale*) – which is awarded by national committees in a centralized evaluation process. The first round of the NSQ took place in 2012 and was followed by two more rounds, in 2013 and 2016-2018.[11]

By achieving the NSQ, scholars gain the mere eligibility for associate and full professorships, while actual hirings still occur at the university department level in a decentralized stage. Coherent with the rationale of the reform, which is to promote research activity and limit local favoritism, the introduction of the NSQ restricts the access to local competitions to candidates whose academic *curriculum vitae* satisfies minimum standards established at the national level. Applicants are evaluated by a committee of five scholars randomly drawn from a list of eligible full professors affiliated with Italian and non-Italian universities. Both the evaluation criteria and the committee composition vary depending on whether a candidate is applying to the NSQ for the associate or the full professorship and on her research field. Academic fields are mapped into 184 different competition sectors, further grouped by the Ministry in two broad macro areas: *bibliometric* (including all disciplines of mathematics, physics, chemistry, earth sciences, health sciences, agronomy and veterinary, engineering and architecture, and psychology) and *non-bibliometric* sectors (humanities, economics, political sciences, and law).

The committee is in charge of screening the items of each candidate's *curriculum vitae* in order to decide whether to award the NSQ or not. The main criterion that committees must follow when evaluating a candidate is her publication record, measured by three observable and well-defined productivity indicators. In *bibliometric* sectors, these indicators are i) the number of articles published in sci-

[10] The hierarchical structure of Italian universities consists of three main ranks: assistant, associate and full professors. Until 2010, the three positions were all tenured and assistant professors were hired under permanent contracts. After the reform, instead, assistant professors are hired under fixed-term contracts.

[11] Differently from the previous two, the 2016-2018 round consists of multiple calls which were opened every four months over a two-years time window.

entific journals, ii) the number of citations received, and iii) the scholar's h-index. In *non-bibliometric* sectors, they are i) the number of monographs, ii) the number of book chapters and articles in scientific journals, and iii) the number of articles published in a list of A-ranked journals. All indicators are calculated over the ten years prior to the NSQ call and normalized by a candidate's academic age. Other criteria include the participation in national or international projects, editorial activities, fellowships, and awards. Although the Ministry allows committees to decide autonomously the weight assigned to each of the aforementioned elements, it also explicitly states that the three productivity indicators should constitute the key criteria.[12]

In particular, the ministry defines specific minimum thresholds for attaining the qualification in each field. These standards are set by looking at the publication records of associate and full professors already employed in the Italian university system. In order to achieve the associate (full) professor qualification in a bibliometric sector, a candidate must score above the median associate (full) professor in her sector **in at least two out of the three** productivity indicators. A similar one-out-of-three rule holds for the non-bibliometric sectors. These rules represent a (almost) necessary but not sufficient condition to achieve the qualification since committee members might deliver a negative judgement even when all of a scholar's indicators surpass the relevant thresholds. Moreover, they also have the right to deviate from the aforementioned rule by awarding the qualification to candidates who do not comply with the productivity requirements. Nonetheless, this latter possibility is allowed only in case of an extremely positive evaluation of the other elements of the *curriculum*.[13]

The first round of the NSQ opened between June and July 2012, when both the call for commissioners and that for candidates were published. In August, the ministry released the sector-specific cutoff values for the each of the three productivity indicators. The deadline for candidates to apply was set for the 20th of November. After this date, the ministry made public candidates' scores, calculated by the ministry, and the list of commissioners in each field, randomly drawn from the list of eligible full professors. Candidates had the right to withdraw their applications until February 2013. This option was particularly important since a

[12]Each committee is composed of four full professors at Italian universities and one at a university in a different OECD country. The eligibility requirements for commissioners are similar to those for candidates: when considering the aforementioned productivity indicators, only above-median full professors can be part of the evaluating committee for a given field.

[13]Figure 1.9 in the Appendix depicts the extent of the deviation from the two-out-of-three rule across different academic fields. The green bars describe the proportion of candidates who obtained the qualification in the 2012 NSQ despite not complying with the two-out-of-three rule. On average, less than 15% of candidates who did not comply with the two-out-of-three rule achieved the qualification in 2012.

negative assessment by the committee in the 2012 round of the NSQ implied that a candidate could not apply to the subsequent one.[14] Thus, scholars could decide whether to undergo the evaluation or not after having observed their scores, the cutoff values, and the composition of the evaluation committee. Most committees completed their work and published the outcome by June 2013, while in few cases, the evaluation process took until December 2013.[15]

1.3 Data and Sample Description

In this study, we combine different administrative and publicly available data sources to build a unique and comprehensive dataset containing, for each candidate for the 2012 NSQ, i) the score in each of the three productivity indicators and the outcome of the qualification procedure; ii) the academic position and affiliation at the time of the call; iii) the complete publication record from 2007 to 2016.

The list of applicants to the 2012 round of the NSQ is obtained from the MIUR website. The administrative records include information about each candidate's application(s), such as the competition sector, the scores in the productivity indicators, the sector-specific cutoffs, and the final outcome of the evaluation procedure. We merge these data with the professor census, which covers all assistant, associate, and full professors employed in the Italian public university system in 2012-2016. This longitudinal database allows us to determine, for each applicant, her position, department of affiliation, and academic field as of 2012, as well as her later promotion patterns. Since the NSQ system allows for multiple applications per candidate, in our baseline specification, we consider each candidate's 'best' application in terms of distance from the cutoffs. However, as shown in Section 1.6.2, results are robust to considering, for each scholar, the outcome of her application in the competition sector in which she was already employed at the time of the call.

Our measures of research productivity come from the *Scopus* database, the largest repository of peer-reviewed literature. We query the *Scopus* archives in order to retrieve each scholar's complete publication record. For each item, we obtain the cover type (article, conference paper, book chapter or review), author's affiliation, publication date, journal name, and the full list of coauthors. Then,

[14] Candidates who received a negative judgement by the committee were prevented from applying for the qualification for the same rank and in the same competition sector for two years, which implies that rejected candidates could not apply to the following one or two rounds, depending on the exact timing of the calls.

[15] Importantly, as also discussed in Section 1.3, our dataset covers all participants to the 2012 call at the time of the deadline (November 2012), thus also including withdrawn applications.

we use this information to build a panel dataset at the scholar-by-year level, including measures of the quantity of publications, the quality of the journals in which they are published, and the citations received. The main journal-specific quality indicator is the 2012 *CiteScore index*, which provides a weighted average of the citations received by each journal in a given year. In order to account for the wide heterogeneity between the different academic fields, we look both at the overall *CiteScore index* and at its within-field counterpart, the *CiteScore journal percentile*. Furthermore, we exploit two alternative measures of journals' prestige: the *SJR* (Scimago Journal Rank) and the SNIP (Source Normalized Impact per Paper).[16] The citations received by each published paper are counted as of July 2017.

Overall, approximately 40 thousand researchers took part in the first round of the National Scientific qualification. We discard applicants in non-bibliometric sectors where the cutoff rule is not strictly enforced by most of the committees.[17] In subjects such as humanities, law, political sciences, and economics, where the number of publications per year is typically lower, the thresholds are often very close – or even equal – to zero.[18] Therefore, since more than 90% of the applicants satisfy the one-out-of-three rule – specifically designed for non-bibliometric sectors – compliance with this rule constitutes a very poor proxy for candidates' quality. Moreover, the resulting lack of observations below the cutoff(s) would not allow us to implement our regression discontinuity design in such fields.

Out of the 20 thousands candidates in bibliometrics sectors – including mathematics, physics, chemistry, earth sciences, health sciences, agronomy and veterinary, engineering and architecture, and psychology – about two thirds apply for the associate professor qualification.[19] Given our focus on promotion incentives, we limit the analysis to the subset of candidates who are already employed as

[16]More precisely, the 2012 *CiteScore* index is computed as the total number of citations received in 2012 by documents published in the three years before, divided by the total number of documents published over the same period. The *CiteScore journal percentile* ranks the journals belonging to each field according to their *CiteScore* index. The *SJR* and *SNIP* indicators are computed in similar way to the *CiteScore index*, thus making them a weighted average of the citations received in a given year by documents published in the three previous years. However, weighting procedures differ from those used to construct the *CiteScore* index.

[17]For 3000 out of the total number of participants in the NSQ it was not possible to identify a unique best application and, consequently, a unique sector. We also exclude them from our analysis.

[18]Given the way cutoffs are established, a threshold equal to zero means that the score of the median associate (or full) professor in that competition sector, for that specific indicator, is equal to zero.

[19]In Section 1.7 we also present our main equation estimated on the sample of associate professors who participated in the NSQ for full professorship in 2012. The process for selecting this sample of applicants to the full professor qualification (4,866) follows the one for candidates to the associate professor qualification.

tenured assistant professors at the time of the deadline. This group (7000 scholars) accounts for about 45% of the total number of applicants to the associate professor qualification. The remaining share of applicants consists of researchers working for non-university institutions in Italy or abroad, academics affiliated with non-Italian universities, and young untenured scholars.[20] Importantly, the wide coverage of *Scopus* allows us to detect a unique author identifier 97% of the candidates in the sample. For the remaining 3% of the scholars, it could be either the case that none of their publications are recorded in the database or that homonymies and misspelled names result in an unsuccessful merge. Lastly, in our baseline specification we discard within-sector outliers, and observations belonging to competition sectors with fewer than 30 applicants thus ending up with a sample of 4920 scholars.[21]

A detailed description of the sample is presented in Table 1.1. A significant proportion of candidates for the associate professor qualification were relatively experienced: the average academic age – that is, the number of years since the first publication recorded in SCOPUS – is approximately 16 years as of 2012. Moreover, they published on average nine papers in the ten years prior to the NSQ. Slightly less than 60% of candidates achieved the NSQ, whereas two-thirds of them satisfied the two-out-of-three rule. Lastly, in the group of academic fields in our sample the number of collaborations is relatively high: only 2% of the papers published by the assistant professors in our sample is single-authored, and the average number of coauthors per publication is 10.82.

Figure 1.1 plots candidates according to the compliance with the two-out-of-three rule – those in the upper-right quadrant, as the figure is drawn for the case when the h-index is below the cutoff – and according to the final outcome of the qualification procedure. The limited degree of fuzziness in Figure 1.1.b confirms that the two-out-of-three rule constitutes a determinant criterion for awarding the qualification in bibliometric fields. Moreover, they also show how the mass of observations concentrates around the multidimensional cutoff, particularly around the intersection of the zero-distance axes. This finding is not surprising since the threshold values are computed by looking at the median associate professor in each field. As discussed in detail in the following section, this particular feature of our data implies that, although local, the effect is estimated in the neighborhood of the representative scholar in each field.

[20] We exclude this group of applicants because we have no information about their employment status in 2012. Thus, we are unable to distinguish between scholars who are employed in Italian universities, although not tenured, from those who work in other institutions and for whom achieving the NSQ does not imply a variation in promotion incentives.

[21] in Section 1.4, we explain how we determine outliers and why we eliminate them; additionally, in Section 1.6.2, we test the robustness of our results to adopting different sample restrictions.

1.4 Empirical Strategy

We exploit the cutoff rule implemented within the NSQ to determine whether a quasi-experimental provision of promotion incentives significantly affects the productivity of a large sample of high-skilled public employees, such as academics. More precisely, in a regression discontinuity framework, we compare the post-call research productivity of barely successful and unsuccessful assistant professors who participated in the 2012 NSQ call. Before describing the details of the empirical methodology, we discuss by what means achieving or missing the qualification exposes candidates to different promotion incentives in the form of different promotion thresholds to be met in the future rounds of the NSQ.

1.4.1 Promotion incentives in the NSQ

The NSQ introduces explicit thresholds that scholars have to meet in order to gain the eligibility for career advancements. Such multiple-step procedure with clear and well-known criteria for promotions therefore entails significant dynamic incentives by making possible, for those ranked lowest, to climb the academic ladder in a few years.[22]

Although earning the qualification does not immediately imply an advancement to the next rung of the career ladder (associate professorship), it sets a new attainable goal to be achieved in the subsequent call of the NSQ: the qualification for full professorship. In order to fulfill this goal, qualified candidates need to enrich their publication record to meet the corresponding bibliometric requirements, that is, the number of articles and citations and the h-index required to achieve the full professor qualification. Figure 1.2 shows that approximately one-third of the assistant professors who successfully took part in the NSQ for associate professor in 2012 were indeed able to also earn the eligibility for a full professorship by 2016.[23]

[22] Although the NSQ might not have been explicitly designed to introduce career-based incentives, as highlighted by Lazear and Gibbs (2014), promotions may constitute an "accidental incentive system" (p. 262). The perspective of career advancement within the organization could enhance employees' motivation and is inextricably linked to the existence of a hierarchical structure.

[23] This share does not necessarily match the one of assistant professors who obtained a chair shortly after 2016, as actual promotion also depends on universities' turnover and budget constraints. However, the decentralized stage is characterized by limited competition – promotions happen mostly within the initial department of affiliation – and qualified candidates move up the academic ranking smoothly: more of two-thirds of successful applicants in 2012 were actually promoted by the end of 2016. We further discuss the importance of the decentralized stage in our setting in Section 1.7

Barely unsuccessful scholars, however, are not exposed to the same high-powered incentives. Although not explicitly ruled out by the institutional setting, in fact, it is extremely unlikely for candidates who do not hold the associate professor NSQ to earn the full professor NSQ. The share of assistant professors who succeeded in the qualification for a full professorship without earning the intermediate step in 2012 is indeed very small (4.5%). Moreover, this number drops to 1% for the 2013 and 2016 rounds. Thus, failure in 2012 implies that a candidate will have to re-apply for the qualification for an associate professorship in the future and, for the marginal unsuccessful applicant, this goal does not require a substantial effort increase since her productivity scores are already very close to the relevant thresholds. Promotion incentives for this subset of candidates are clearly weaker or even absent.[24]

In principle, those scholars who have to postpone their career progression could be forward looking, already targeting the requirements for the full professor qualification. However, they face more uncertainty, as they have to rely upon the stability of the institutional setting over a longer horizon. Additionally, since the ministry sets the minimum thresholds using the median scholar as the reference point, a new inflow of full professors is likely to affect the productivity distribution and, consequently, the future realization of the cutoffs. These two sources of uncertainty weaken the incentives to target the eligibility requirements for the full professor qualification before succeeding in the associate professor one.

1.4.2 A triple (fuzzy) regression discontinuity design

Our regression discontinuity strategy exploits the discontinuous jump in the probability of obtaining the qualification, arising when two of the three indicators crosses its relevant threshold. By fully modeling the two-out-of-three rule with three forcing variables – the productivity indicators – we are able to define a three-dimensional cutoff, that is, a hyperplane that is the $\mathbb{R}^3$ equivalent of the standard single-variable frontier. Therefore, one should picture the discontinuity in the probability of receiving the NSQ around the 3-D frontier as a pooled or combined version of the smaller, single-variable, discontinuities. Since the compliance with the two-out-of-three rule alone does not represent a sufficient condition to achieve the qualification, the probability of receiving the treatment will jump by less than 100% when crossing the multidimensional cutoff. Hence, our empirical strategy relies on a (triple) fuzzy regression discontinuity design with multiple sector-specific cutoffs.

[24]Candidates who withdraw their applications after observing their scores and relevant cutoffs can re-apply in the first subsequent round of the NSQ. Rejected candidates, instead, must wait at least two years to apply again for the associate professor qualification in the same competition sector.

Formally, let us define the assignment variables – number of articles, number of citations and h-index – as x_{i1}, x_{i2} and x_{i3}, respectively. Then, G_{iks} is an indicator function that equals one when score k of candidate i belonging to competition sector s is strictly above the cutoff m, that is

$$G_{iks} = \begin{cases} 0 \text{ if } x_{iks} \leq m_{ks} \\ 1 \text{ if } x_{iks} > m_{ks} \end{cases} \quad \text{for each } k \in \{1, 2, 3\}.$$

The indicator D_{is} thus describes the aforementioned two-out-of-three rule:

$$D_{is} = \begin{cases} 0 \text{ if } \sum_{k=1}^{3} G_{iks} < 2 \\ 1 \text{ if } \sum_{k=1}^{3} G_{iks} \geq 2. \end{cases}$$

Consequently, our first-stage equation is

$$Q_{is} = \alpha_0 + \alpha_1 D_{is} + f(x_{iks} - m_{ks}) + \alpha_2 Z_s + \nu_{iks}, \tag{1.1}$$

where Q_{is} is an indicator that equals one when a candidate achieves the qualification, $f(x_{iks} - m_{ks})$ is a flexible nonlinear function of the distance of the running variables from the threshold(s) (including 2nd order polynomials of the three variables and all possible interactions), and Z_s are sector-specific fixed effects. Analogously to a 'canonical' RD design – with a single running variable and single cutoff – the coefficient α_1 measures the discontinuous jump in the probability of achieving the qualification that arises when a candidate complies with the cutoff rule. More precisely, α_1 captures a weighted average of the discontinuity in the probability of achieving the qualification when crossing the frontier hyperplane from all the octants not satisfying the two-out-of-three rule.

This discontinuity in the probability of obtaining the qualification is then used as an instrumental variable to estimate our second-stage equation, which is

$$Y_{is} = \gamma_0 + \gamma_1 \hat{Q}_{is} + f(x_{iks} - m_{ks}) + \gamma_2 Z_s + \eta_{iks}, \tag{1.2}$$

where γ_1 is the local average treatment effect (LATE) of achieving the NSQ in 2012 on any of our measures of scientific production Y_{is}, computed in the post-call period. The corresponding reduced form equation is

$$Y_{is} = \beta_0 + \beta_1 D_{is} + f(x_{iks} - m_{ks}) + \beta_2 Z_s + \eta_{iks}, \tag{1.3}$$

where β_1 measures the intention-to-treat (ITT) effect of complying with the two-out-of-three rule. The interpretation of γ_1 and β_1 in this multidimensional regression discontinuity framework is analogous to that provided for the α_1 coefficient of the first stage: they capture a weighted average of the effect of crossing the 3-D frontier from all the neighboring octants.

To account for the heterogeneity between different academic fields, we allow $f(\cdot)$ to be fully flexible across sectors in Equations 1.1, 1.2 and 1.3 by interacting each assignment variable – centered around its sector-specific cutoffs –, their squared values, and their first and second degree interactions, with the competition sector dummies.[25]

Because of both the complexity of the framework and the lack of a standard procedure to compute joint bandwidths in a multidimensional regression discontinuity design with multiple cutoffs, our preferred specification is a fully-parametric one in which, to reduce the weight of potential outliers, we exclude candidates in the top decile or in the bottom percentile of the distribution of the distance from the cutoff. Then, in Section 1.6.2, we show that results are robust to considering a linear specification within an arbitrary range around the zero-distance cutoff(s) and to adopting alternative sample restrictions to deal with outliers.

Finally, it is important to stress that our identification strategy is less vulnerable to the main criticism usually made for regression discontinuity designs, namely, the locality of the estimated effect. The estimated discontinuity is indeed a weighted average of the discontinuities along the three different frontiers, one for each productivity indicator. Furthermore, the marginal candidate in this setting is a representative scholar in her field since cutoff values are set by looking at the median professor in each competition sector. As a result, a large mass of observations is concentrated around the three-dimensional frontier, as highlighted in Figure 1.1.

1.4.3 Validity of the RD design

Our identification strategy relies on two main assumptions: 1) the probability of achieving the qualification jumps discontinuously at the multidimensional cutoff describing the two-out-of-three rule; and 2) the joint distribution of the running variables does not exhibit any bump immediately in the neighborhood of the same cutoff. Furthermore, in a full-parametric, multidimensional regression discontinuity design, special attention should be devoted to possible misspecification issues (3). While the satisfaction of Assumption 1 is discussed in Section 1.5, we address 2 and 3 here.

Testing the validity of Assumption 2 is crucial to discarding two potential threats for our identification strategy: manipulation and sample selection. Re-

[25]The whole evaluation process should be seen indeed as a combination of many small selections, since candidates face different thresholds and committees depending on the competition sector for which they applied. Figure 1.8 in the Appendix shows the extent of such across-field heterogeneity in the cutoff values: in many subfields in medicine and physics the median number of articles among associate professors, over the 2002-2012 period, is above 40 articles, while it is often below 10 in mathematics and engineering.

garding the former, the possibility for candidates to manipulate their publication records in order to meet the minimum standards seems remote since both individual scores and thresholds are computed by the Ministry of Education, University and Research. The Ministry collects candidates' full publication records from their application webpage and cross-validates each research item by querying the two largest databases of peer-reviewed literature: *Scopus* and *Web of Science*.[26] Moreover, because of the short time frame between the publication of the call and the application deadline, it is unlikely that scholars would have the time to adjust their publication records to meet the established requirements. Regarding selection, a positive jump in the density could also reveal that scholars who decide to participate in the NSQ without complying with the two-out-three-rule constitute a selected sample. For instance, one potential concern could be that scholars below the cutoff were disincentivized from applying given that a negative evaluation by the committee would have prevented them from participating in the subsequent round of the NSQ. However, it is important to remark that candidates are given the opportunity to withdraw their application after having observed their precise scores and the composition of the committee, and prior to the evaluation itself (that is, by February 2013). Hence, applying to the 2012 NSQ is relatively costless, even for those below the thresholds, and selection concerns should be limited as our sample of candidates is based on the list of applications at the time of the deadline (November 2012), thus including withdrawn applications.

We formally test whether the distribution of candidates is discontinuous around the cutoff. In Figure 1.3, we report the frequencies, as well as the density and confidence intervals computed following McCrary (2008), for each of the three forcing variables centered around the cutoff. None of the running variables exhibits a significant (at the 10% level) discontinuous jump in the density in the neighborhood of the zero-distance from the cutoff. Point estimates (standard errors) of the density test are 0.085 (0.077) for the number of articles, 0.083 (0.075) for the number of citations and 0.043 (0.046) for the h-index. Furthermore, Figure 1.10 in the Appendix provides additional support for the assumption that scholars do not endogenously sort or select around the threshold. The robust bias-corrected manipulation test proposed in Cattaneo et al. (2017b) delivers non-significant estimates for each of the three running variables (number of articles: T=-0.79, p-val=0.43; number of citations T=-0.27, p-val=0.78; h-index: T=-0.96, p-val.=0.33).[27]

Finally, our fuzzy regression discontinuity design also relies on the assumption that scoring above the median professor in two out of the three bibliometric indicators should have no impact on future scientific productivity other than that

[26] More precisely, a ministerial agency (ANVUR) computes both the individual scores and the thresholds.

[27] All estimates are obtained using the Stata package described in Cattaneo et al. (2017a).

passing through the achievement of the NSQ. It seems, however, extremely unrealistic that other confounding factors or policies could drive the observed jumps at such particular cutoffs.

1.5 First Stage

A crucial condition must hold to implement our empirical strategy: overcoming the bibliometric thresholds and satisfying the two-out-of-three rule must result in a discrete jump in candidates' probability of achieving the qualification. In this section, we show that this is indeed the case. In Table 1.3, we report both the estimates of the first-stage equation when considering each of the three bibliometric indicators and its corresponding cutoff, separately (Columns 1 to 6), and when exploiting the three running variables simultaneously, as formalized by Equation 1.1 (Column 7). The estimated coefficient from this triple-RD – that is, our preferred specification – shows that compliance with the bibliometric two-out-of-three rule discontinuously increases the probability of achieving the qualification for an associate professorship by approximately 30 percentage points. The magnitude of the first stage confirms that commissioners attribute a strong weight to the compliance with the two-out-of-three rule when making their decisions.[28]

The single-RD estimates are also positive and significant in all specifications, consistent with the graphical evidence in Figure 1.4. In Columns (1), (3), and (6), we estimate the discontinuity in the probability of achieving the qualification when passing each of the three bibliometric threshold – the number of articles, the citations, and the h-index – , assuming a quadratic functional form on the entire support and including both academic field fixed effects and field-specific interactions. In Columns (2), (4), and (6), we replicate the same estimates assuming a linear functional form within the MSE-optimal bandwidths. In this case, to take into account the wide between-field heterogeneity in candidates' average productivity, we use as running variables the relative distances from each threshold, that is, the original running variable divided by the threshold itself. By doing so, we are also able to compute three optimal bandwidths, expressed in relative terms, which can be used across the different fields.[29] The estimation results are very close to their fully parametric counterparts. Of course, the magnitude of each single-RD coefficient is lower than that resulting from the triple-RD estimation since the former measures the discontinuous jumps around each single threshold regardless of whether the specific indicator is pivotal for the compliance with the

[28]The corresponding estimates for the sample of candidates to the NSQ for full professorship are presented in Table 1.10 in the Appendix.

[29]Specifically, the optimal-MSE bandwidths are computed following Calonico et al. (2014) for each of the three relative distances, separately.

two-out-of-three rule. Hence, estimating three standard, single-forcing variables RDD would not account for the compliance (or defiance) with the other two requirements, thus increasing the degree of fuzziness. This is precisely the reason why we adopt a triple-RD design, in which the α_1 coefficient of Equation 1.1 should be interpreted as a combined version of three smaller discontinuities.

1.6 Results

1.6.1 Quantity of publications

Table 1.4 reports the main result of our empirical analysis: achieving the qualification for an associate professorship in 2012 – thus being provided with higher promotion incentives – has a positive impact on the number of papers published in the subsequent years. The local average treatment effect (LATE) of achieving the qualification on the number of scientific publications over the 2013-2016 period corresponds to 6.5 publications and is 3.25 times larger than the intention-to-treat (ITT) effect of complying with the two-out-of-three rule (which is equal to 2 publications). Both the LATE and the ITT coefficients are statistically significant at the 1% level. The estimated effect is sizable and corresponds to approximately 40% (LATE) of the sample average number of publications over the same period. By looking at the different publication types, we find the effect to be driven mostly by an increment in the number of published articles and, to a smaller extent, reviews and conference papers (see Table 1.4).

In principle, the estimated effect could be due not only to the increased productivity of barely qualified scholars but also to a decline in publications by narrowly unsuccessful candidates. This latter group of scholars might indeed become frustrated and discouraged or could revise their research production function after missing the qualification. In order to disentangle these two hypotheses – the discontinuity being driven by marginal successful or unsuccessful applicants – we exploit the panel dimension of our dataset and replicate our baseline estimation using the yearly number of publications before and after the first call of the NSQ as the dependent variable of interest.

Figure 1.5 reports the estimated LATE of the qualification on the number of publications for each year between 2007 and 2016. Blue diamonds describe the evolution of the number of publications for candidates who marginally missed the qualification in 2012, while red circles depict the corresponding trend for those who barely met it. The vertical distances between the two trends represent the estimated discontinuity in each year (the estimated coefficients are reported in Table 1.13 in the Appendix). The annual productivity of narrowly unsuccessful scholars remains constant in the post-NSQ period, while that of barely successful

applicants exhibits a significant rise. This effect is persistent, large in magnitude, and significant for the whole post-call period, with the only exception of 2015 when the discontinuity is still positive but the larger variance in the data lowers its significance.

The estimates in Figure 1.5 and Table 1.13 also lend strong support to our identification strategy. For the entire pre-NSQ period, the difference between treated and non-treated individuals is close to a precise zero. Hence, the results are not driven by any *ex ante* difference between candidates or by a possible misspecification of the functional form assumed when estimating the relation between the treatment and outcomes.

Additionally, we investigate whether the effect of passing the NSQ in 2012 is heterogeneous depending on candidates' gender and academic field. The estimates are presented in Table 1.5 and 1.14, respectively. When looking at gender heterogeneity, we find the LATE to be homogeneous across female and male candidates: the promotion incentives associated with the achievement of the qualification are equally effective, regardless of gender. However, we find a negative and significant coefficient for the interaction between the female indicator and the one for compliance with the bibliometric rule when estimating the first-stage equation. Hence, women who satisfy the two-out-of-three rule are less likely to achieve the qualification than men with comparable publication records. This result is consistent with the evidence provided by Bagues et al. (2017) and De Paola and Scoppa (2015) – who also document that female candidates have lower success rates in the Italian qualification procedure – and could be due to gender discrimination.

We find moderate evidence of between-subject heterogeneity. The interaction coefficients of the field-specific dummies with the treatment are positive (with the only exception of psychology) even though heterogenous in magnitude and not always statistically different from zero, suggesting that the average effect is not driven by the behavior of scholars belonging to a few peculiar fields.

1.6.2 Robustness checks

Our triple regression discontinuity model is an extended version of the regression discontinuity with multiple assignment variable proposed, among others, by Papay et al. (2011) and Papay et al. (2014). In particular, it is close to what the latter define as the 'Response-Surface RD'. These models depend heavily on a correct specification of the parametric functional form, as the gain in both efficiency and power resulting from multidimensionality comes at the expenses of lower flexibility.[30] Moreover, as for any full-parametric approach, the presence of

[30]Since we want to estimate the average treatment effect along the multidimensional borders, we cannot include a two- or three-dimensional spline since, by doing so, we would estimate a very

(within-sector) outliers can bias the estimated coefficients, as all observations are assigned an equal weight irrespective of their distance from the cutoff.

To account for this latter issue, in our baseline specification, we exclude observations in the top decile and the bottom percentile of the distribution of distances from the cutoffs. In this section, we show that our main results are robust to adopting alternative sample restrictions. More specifically, we replicate our analysis varying the lower and the upper bounds of the distribution of distances from the cutoffs, thus progressively excluding candidates whose scores lie outside specified inter-percentile ranges (see Figure 1.11 in the Appendix). Results from this test show that considering a broader or a narrower sample does not deliver very different estimates for the ITT effect. However, including observations in the far right tail of the productivity indicator distribution increases the noise in the sample and lowers the significance of the estimated coefficients. This finding is consistent with the fact that a candidate's publications have in principle no upper limit, whereas they cannot be less than zero. Thus, most of the outliers are located above the multidimensional cutoff.

To address possible concerns owing to the functional form assumed in our baseline estimation, we replicate our analysis here by assuming a linear specification in the neighborhood of the thresholds. More precisely, we first normalize each running variable through dividing it by the corresponding cutoff value – thus accounting for between-field heterogeneity in candidates' average productivity – and then select three different bandwidths, one for each running variable. Finally, we re-estimate Equation (2) on the sample of scholars whose productivity indicators lie within the resulting multidimensional joint bandwidth, assuming a linear specification.[31] Table 1.11 in the Appendix reports the result of this further robustness check and a comparison with our baseline results. The point estimates resulting from this local linear approach are very close in magnitude to those obtained assuming a second-degree polynomial form over the entire support. However, they are less precise, as standard errors are larger. Our preferred, fully parametric specification with field-specific interactions indeed allows us to better estimate the effect of complying with the two-out-of-three rule accounting for the between-field heterogeneity in the distribution of the productivity indicators. This goal is harder to achieve with a nonparametric approach within the neighborhood of the thresholds since we face a framework with multiple running variables and multiple field-specific cutoffs. The literature indeed lacks a procedure to compute optimal bandwidths in a similar context while at the same time accounting for the wide across-field heterogeneity in the distribution of the running variables.

local effect at the intersection of all cutoffs.

[31]The bandwidths for the three productivity indicator are the MSE-optimal bandwidths computed separately for each running variable, following Calonico et al. (2014)

Additionally, we perform two placebo exercises to rule out the concern that our findings could be driven by systematic differences between candidates at the two sides of the cutoff rather than by a reaction to the treatment provision. First, we estimate our equations using the quantity of candidates' publications in each year before the 2012 NSQ as the dependent variable. In the case of any specification or sorting issues, our triple regression discontinuity model should also deliver non-zero results in the pre-treatment period. As shown in Table 1.13 in the Appendix, no discontinuity in terms of total publications and articles between treated and controls emerges when looking at each of the four years prior to the NSQ. Moreover, in Table 1.2, we show that the marginal applicants in the two sides of the cutoff do not differ in terms of the aggregate quantity and quality of their publications or in the number of collaborations when these measures are computed over the whole 2009-2012 period. Second, we apply a perturbation to each field-specific threshold. We expect the magnitude of both our estimated first-stage and ITT coefficients to decline and the associate confidence intervals to broaden the farther we advance from the original cutoff(s). More specifically, we reshuffle the cutoff values by adding a randomly generated error component $\epsilon \sim N(0, \sigma)$, which is defined as a percentage of the original field-specific cutoff.[32] The resulting perturbation, which we impose to lie within plus and minus the 100% of the original cutoff value, then has a different intensity depending on the standard deviation (σ) of the error. We then estimate the LATE from our baseline regression for increasing values of σ, replicating this exercise for 30 different draws from the ϵ distribution. We show in Figure 1.12 that the magnitude of the effect is the highest in the zero-perturbation case – that is, when using the true threshold values – and decreases in the variance of the perturbation. Taken together, the results from these two robustness tests confirm that our findings are not driven by any *ex ante* difference between candidates at the two sides of the multidimensional cutoff, thus lending important support to our identification strategy.

As a last robustness check, we test whether our results hold when using a different approach to deal with multiple applications. Since the rules of the NSQ allow candidates to apply for the qualification in different competition sectors, in our baseline specification, we consider for each candidate her 'best' application, that is, the one in which she scores the highest in terms of distance from the relevant thresholds. Here, we replicate our analysis considering for each applicant the indicators, the cutoffs, and the qualification outcome in the competition sector to which she already belongs as an assistant professor at the time of the application. Table 1.12 shows that the effect of achieving the qualification on the number of

[32] We first generate the error $\epsilon \sim N(0, \sigma)$ and then draw from the ϵ distribution in order to assign a different perturbation to the cutoff value of each sector. We do this to account for the between-sector heterogeneity in each of the three productivity indicators.

articles published between 2013 and 2016 is still positive and significant under this alternative specification. Coefficients are slightly lower in magnitude, consistent with the fact that, in this case, barely unsuccessful candidates might have succeeded in another competition sector. Therefore, a significant share of the candidates below the multidimensional cutoff are actually qualified and consequently exposed to promotion incentives, which makes the discontinuity in terms of post-call productivity smaller.

1.6.3 Additional results

After analyzing the impact of passing the NSQ on the quantity of published items, we explore in this section whether it also affects other dimensions of the research activity of the academics in our sample. In particular, we look at whether any significant discontinuity between (barely) qualified and non-qualified candidates emerges in terms of citations, publication quality and academic network size.

Citations. By replicating our baseline specification using the post-2012 citations received by each scholar as the dependent variable, we find that passing the qualification for the associate professorship also affects scholars' citations of their work. The results in Table 1.6 show that for papers published from 2013 to 2016, barely successful candidates receive on average 44 citations more than their barely unsuccessful colleagues (Column 1). This result can be attributed largely to both the increased number of publications of qualified scholars and to the increase in the average number of citations *per* paper (Column 2). The probability of publishing an article with more than 50 citations (Column 3) or a non-cited article (Column 4) does not exhibit any jump, however.

Thus, scholars who are provided with higher promotion incentives in 2012 not only increase their publications but also manage to improve on another dimension that is taken into account in the qualification procedure: the number of citations received. This effect is in part simply driven by the increased research productivity of qualified scholars but could also reflect an augmented effort to promote and disseminate scientific works, greater visibility following a promotion, or an increase in the average publication quality. This latter aspect seems of particular importance and is therefore the next dimension on which we focus.

Average publication quality. The publication quality does not directly enter among the productivity indicators considered in the NSQ but could be indirectly affected by qualified scholars' incentives to maximize both citations and publications in a direction that is *a priori* ambiguous. On the one hand, publishing in better, more prestigious journals can increase a scholar's citations and h-Index. On the other hand, there is a potential tension between the quantity and quality of publications, as submissions to prestigious journals are costly, especially in in terms of time, owing to the higher standards required and the more selective

review processes. This trade-off could induce qualified scholars to sacrifice the quality dimension in order to minimize publication times and quickly increase their publication records.

We test these hypothesis by replicating our analysis using as dependent variables several alternative measures of a journal's quality and prestige. Specifically, we consider the *CiteScore*, the *Sjr* and *Snip* indexes, and the within-field *CiteScore* ranking – that is, a measure grouping journals according to their position in the field-specific distribution of the *CiteScore* index. According to the results reported in Table 1.7, Columns (2) to (5), no significant discontinuity in the average publication quality emerges, as all coefficients are not statistically different from zero. Additionally, we test whether the probability of publishing in a journal ranking in the top percentile of the CiteScore index (Column 1) or in a journal with no available measures of quality in the *Scopus* database (Column 6) changes discontinuously at the multidimensional threshold and find that this is not the case.

Hence, the documented increase in publications and citations by barely qualified scholars is not associated with a contemporaneous change in their average publication quality. Importantly, the large increase in the number of publications induced by the provision of promotion incentives does not appear to come at the expense of the average quality.

Co-author network. Finally, we study whether the outcome of the 2012 NSQ has any effect on the number of collaborations or on the size of scholars' co-authors network. In Table 1.8, we report the estimated coefficients from our ITT and LATE equations, using as dependent variables i) the mean and the median number of authors per paper, ii) the probability of publishing a single-authored paper, and iii) the number of distinct co-authors. The three variables are computed for the 2013-2016 period. While the first two outcomes measure how each research paper is produced – that is, whether scholars tend to publish more or less coauthored works – the third proxies for the size of the academic network. We find suggestive evidence of a positive effect of achieving the associate professor qualification on scholars' co-authoring decisions, although the only significant (at the 10% level) coefficient is that for the median number of co-authors. Specifically, the estimated LATE in Column (2) shows that the median paper published by a barely qualified scholar has 2.2 more coauthors than that published by a barely unsuccessful scholar. Thus, these findings suggest that although some scholars might strategically expand their academic network in order to meet the thresholds for the full professor qualification more quickly, this behavior does not seem to be the main driver of the increase in productivity documented in the previous sections.

1.7 Distributional Effects and Competing Mechanisms

1.7.1 The distributional effects of promotion incentives

In this section, we dig deeper into the promotion incentives mechanism and explore whether scholars' reaction differs depending on the intensity of the incentives. Thus, we exploit across-individual differences in the distance between the productivity indicator and the full professor thresholds in 2012, the latter being the best estimate a candidate can have about the future thresholds she will face. This distance measures the size of the gap a scholar needs to fill in order to pass the (future) full professor threshold and therefore proxies for her chances of meeting the promotion thresholds in a relatively short time interval. Thus, we expect incentives to be low when the probability of obtaining the full professor qualification in the short or middle run is close to zero or to one, that is, when the gap that scholars have to fill is either too large or too small (see (Lazear and Gibbs, 2014)).

To reduce the dimensionality of the problem – and consistently with the dependent variable in our main regression (the number of publications post-2012) – we focus on the distance between the first bibliometric indicator – the number of articles published over the 2002-2012 period – and its (field-specific) full-professor cutoffs. Moreover, in order to account for the heterogeneity in pre-2012 research productivity and field characteristics, we normalize this distance dividing it by each candidate's number of publications as of 2012. The resulting index therefore measures the relative increase in publications that a candidate has to produce in order to reach the full professor cutoff. Thus, it constitutes a measure of the attainability of a further promotion in one of the subsequent calls of the NSQ.[33]

Figure 1.6.a reports point estimates and the associated 95% confidence intervals from regressing the number of post-2012 publications on our treatment, interacted with a categorical variable grouping observations according to the above-defined index. The estimated effects of the interaction terms show that the relationship between the increase in productivity and the distance from the promotion threshold is inverted-U shaped. The intention-to-treat effect is not significantly different from zero for those assistant professors who, in 2012, were either too close or too far from meeting the requirements for a full professor qualification. On the contrary, it is strongest for those in the middle of the relative-distance distribution. In order to fill the gap with the full professor cutoff, scholars in the third quartile would have to increase their stock of publications by approximately

[33] More precisely, the index is defined as $dist_{i,1,s} = \frac{m_{1,s}^{full} - x_{i,1}}{x_{i,1}}$, where $m_{1,s}^{full}$ is the field-specific cutoff to overcome – in terms of the number of articles – to achieve the full professor qualification, while $x_{i,1}$ the professor's score in the same indicator.

25%.[34] This goal is realistic in a short- or middle-run horizon. Conversely, scholars in the last quintile would need to almost triple their stocks of publications, a target that is much more difficult to meet in a relatively short time interval. Importantly, this heterogeneity in the effect is not driven by across-quintiles differences in the probability of achieving the associate professor qualification in 2012 when complying with the two-out-of-three rule or in the pre-2012 research productivity of candidates. Both the first-stage coefficient and the ITT coefficient in the pre-2012 regression are indeed stable across the different quintiles (see Figure 1.13 in the Appendix). Furthermore, and consistent with the described mechanism, Figure 1.6.b documents that the probability of actually achieving the qualification for a full professorship by 2016 is heterogeneous across the above-defined quintiles. Candidates who were already very close to the full professor cutoff and those who increased their publication records the most after achieving the qualification are also those who are more likely to effectively achieve the qualification for a full professorship by the end of 2016. Conversely, candidates in the last two quintiles have a much lower likelihood to succeed in the full professor NSQ in one of the following rounds.

Taken together, these two pieces of evidence show that candidates provided with the strongest incentives are also those who increase their post-2012 productivity the most, thus effectively improving their chances to succeed in one of the subsequent full professor qualification procedures. The effort induced by promotion incentives translates into an effectively higher probability of success. These results not only shed light on the distributional consequences of the promotion incentives induced by the qualification process but also lend important support for the promotion incentives mechanism. Most of the alternative explanation for our results – such as qualified scholars obtaining different teaching duties or easier access to research funds – would clash with the observed heterogeneity of the effect, depending on the variation in the intensity of the promotion incentives.

1.7.2 Promotion incentives *vs.* competing mechanisms

Results from our analysis document that scholars who attain the NSQ in 2012 increase the quantity of publications in the four years following the call. We interpret this finding as a response to the provision of promotion incentives: gaining the eligibility for an associate professorship 'unlocks' the possibility to achieve the qualification for a full professorship and incentivizes scholars to meet the requirements for the full professor qualification in the subsequent round.

Certainly, there might be competing explanations for our result. For instance, if perceived as a reward for past effort, obtaining the qualification could have a

[34]The stock of publications is computed in the ten years prior to the NSQ.

motivational effect, thus enhancing productivity. Moreover, passing the qualification could induce substantial changes in scholars' daily life, as career advancements in academia are possibly associated with different teaching or bureaucratic duties, better access to research funds or broader networks. Still, this latter hypothesis seems inconsistent with the observed timing of the effect. Scholars' productivity begins rising immediately after the attainment of the mere eligibility for an associate professorship, rather than at the time of the actual promotion, which, for more than 75% of the qualified candidates, did not take place earlier than 2015. As a further exercise to address these concerns, we estimate our baseline equation for the sample of associate professors who apply for the full professor qualification in 2012. The NSQ indeed regulates both the access to associate and full professor positions, but candidates for this latter rank will have vanishing career incentives once the goal is achieved since no further advancements are possible. As a result, any effect detected in this sample of participant can hardly be reconciled with the promotion incentive mechanism proposed in this study.

The estimates reported in Table 1.9 show that applicants who barely earn or barely miss the eligibility for the top academic position do not exhibit any significant difference in terms of later research productivity. This zero or even negative effect clashes with several alternative explanations for our main result. It shows that the effect of achieving the qualification is specific to the group of academics (tenured assistant professors) with further career prospects.

We also test whether the observed increase in publication might be due to the competition at the decentralized stage, where associate professorships are actually awarded. Achieving the NSQ might indeed incentivize qualified scholars to publish more in order to maximize their chances of obtaining an associate professor position as soon as a job vacancy opens rather than to meet the future full professor thresholds. Although data on scholars' promotion patterns suggest that there is limited across- and within-department competition, it could still be the case that the productivity jump is driven by fields with few vacancies and many qualified scholars competing for a position.[35]

Hence, we exploit the across-sector heterogeneity in the degree of internal competition for being hired as or promoted to an associate professor and test whether the effect of promotion incentives is actually stronger in sectors that feature more competition at the decentralized stage. Since we do not observe the actual number of vacancies but rather the equilibrium outcome, we use the ratio between the number assistant professors with an associate professor qualification and the number of existing associate professors in each field at the end of 2012

[35] Approximately two-thirds of eligible candidates in our sample obtained an associate professorship within three years from achieving the NSQ. Additionally, 97% of them obtained a promotion within the same university where they were employed in 2012.

as a proxy for the degree of competition at the academic field level. This ratio indeed measures the ease of access to an associate professor position in a given competition sector, conditional on having achieved the qualification. Sectors in which there is a large mass of qualified candidates and few associate professors on staff are indeed likely to have lower turnover rates and therefore fewer vacancies, which make them relatively more competitive than those with a relatively low share of qualified scholars. Figure 1.7 shows that the effect of passing the NSQ on later productivity (number of publications) is very homogeneous across competition quintiles. This evidence, together with the distributional effects discussed in the previous section and with the zero-effect found on the sample of candidates to the full professor NSQ, strongly support promotion incentives as the main mechanism at work.

1.8 Conclusion

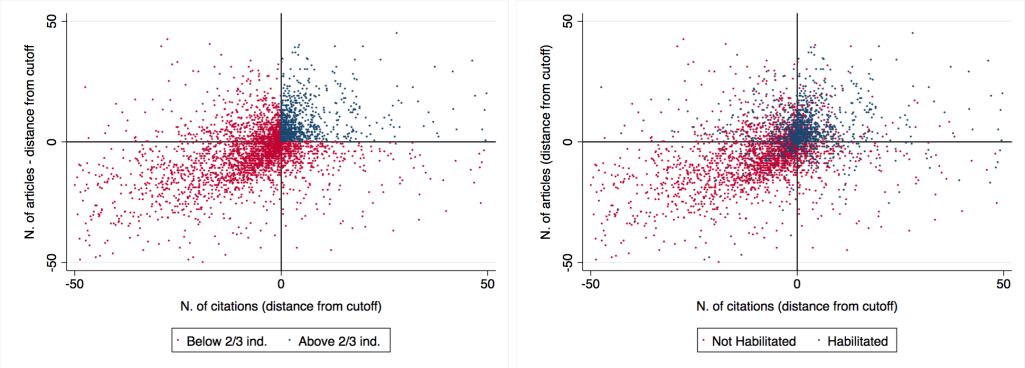
This paper studies the effectiveness of promotion incentives for high-skilled public employees: academics. For a sample of 5,000 tenured assistant professors participating to the first round of the Italian NSQ – the centralized evaluation procedure awarding the eligibility for career advancements – we find that scholars exposed to a quasi-random increase in promotion incentives in 2012 owing to a success in the NSQ increase the number of publications by almost 40% over the 2013-2016 period. Additionally, we find that the effect of promotion incentives is strongest for those scholars who are neither too far nor too close from the relevant future promotion thresholds. That is, promotion incentives are most effective when the promotion is "neither too hard to achieve, nor too easy" (Lazear and Gibbs, 2014, p.269). When exploring additional aspects of scholars' research activity, we find that qualified candidates receive more citations after the achievement of the NSQ and tend to expand the number of collaborations. The average publication quality – proxied by several measures of the journal's prestige – remains constant.

Several robustness tests and placebo exercises confirm the validity of our (triple) fuzzy regression discontinuity design. Further, consistent with the identification hypothesis according to which achieving the qualification affects productivity only through an increase in promotion incentives, we do not find a similar effect in the sample of associate professors applying for a full professorship. Once the top ladder of the academic hierarchy is reached, achieving the qualification does not provide any further promotion incentives. Finally, we provide evidence that our results are not driven by possible changes in scholars' routine associated with a promotion or by the competition at the decentralized stage among qualified candidates.

These results shed light on an important and relatively unexplored topic in the existing personnel economic literature: the efficacy of promotion incentives in the public sector, where performance-pay schemes are typically difficult to implement. Our findings show that promotion incentives, in the form of well-defined promotion thresholds, can effectively enhance public workers' productivity, especially when meeting the established targets requires a substantial but not excessive provision of effort. This conclusion is particularly important for all of those countries in which the state personnel represents a significant share of the overall workforce – such as most European countries – and therefore, enhancing their performances can foster the country's overall productivity.

1.9 Figures and Tables

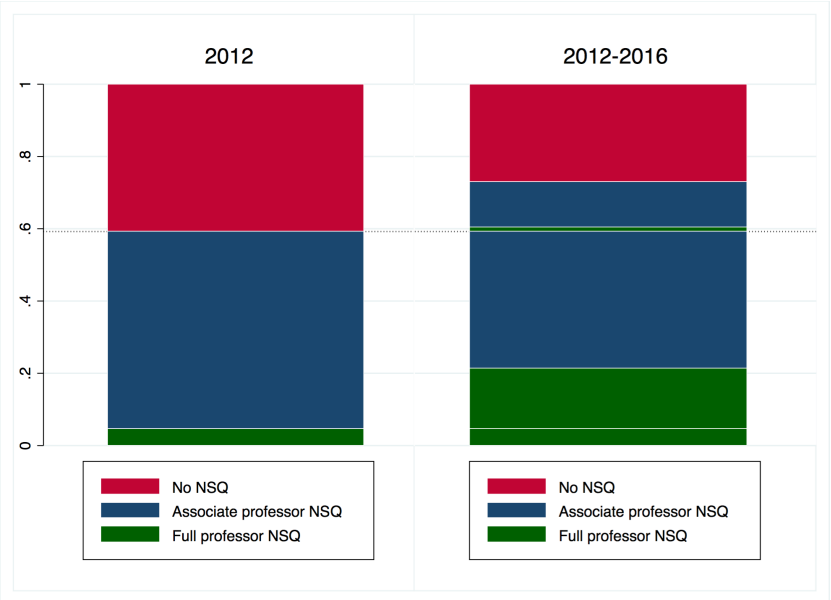
Figure 1.1: Distribution of candidates with respect to distance from cutoffs



a. Compliance with the two-out-of-three rule **b.** Qualified

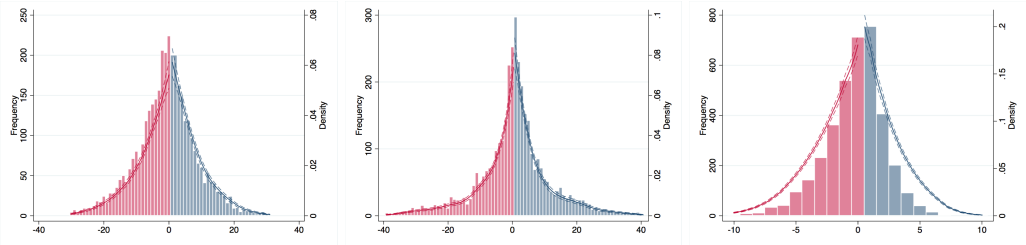
NOTES. This figure depicts the distribution of candidates for an associate professorship in the 2012 NSQ in bibliometric fields, depending on the distances between the first two bibliometric indicators and the corresponding cutoffs. The distance between the number of articles and the cutoff is on the y-axis, while the distance between the number of citations and the cutoff is on the x-axis.In both panels, we consider only applicants whose H-Index is below the cutoff. Therefore, circles in the upper-right quadrant correspond to candidates complying with the two-out-of-three rule. Blue circles indicate qualified candidates, while red circles indicate unsuccessful candidates. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the running variables. We also exclude fields with more than 90% of successful candidates and those with less than 30 observations.

Figure 1.2: NSQ trajectories 2012-2016



NOTES. This figure depicts the share of successful and unsuccessful assistant professors applying to the NSQ between 2012 and 2016 in bibliometric fields. The left bar reports the share of candidates who were not qualified (red), qualified for an associate professorship (blue) and qualified directly for a full professorship (green). For each of the three groups, in the right bar we report the corresponding shares as of the end of 2016, that is, after the 2013 and 2016 calls of the NSQ.

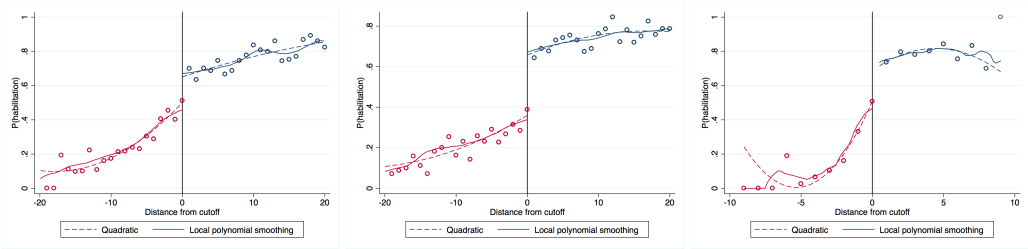
Figure 1.3: Frequency distribution and manipulation test



a. N. of articles **b.** N. of citations **c.** H-index

NOTES. This figure depicts the frequency distribution and the kernel density estimation of candidates for an associate professorship in the 2012 NSQ in bibliometric fields depending on their distance from each of the three bibliometric cutoffs. The frequency distributions in Panels A, B, and C are constructed within the interval $[-30, 30]$ $[-40, 40]$ $[-10, 10]$, respectively. In all panels, the bin width is equal to 1. The kernel density is estimated following McCrary (2008).

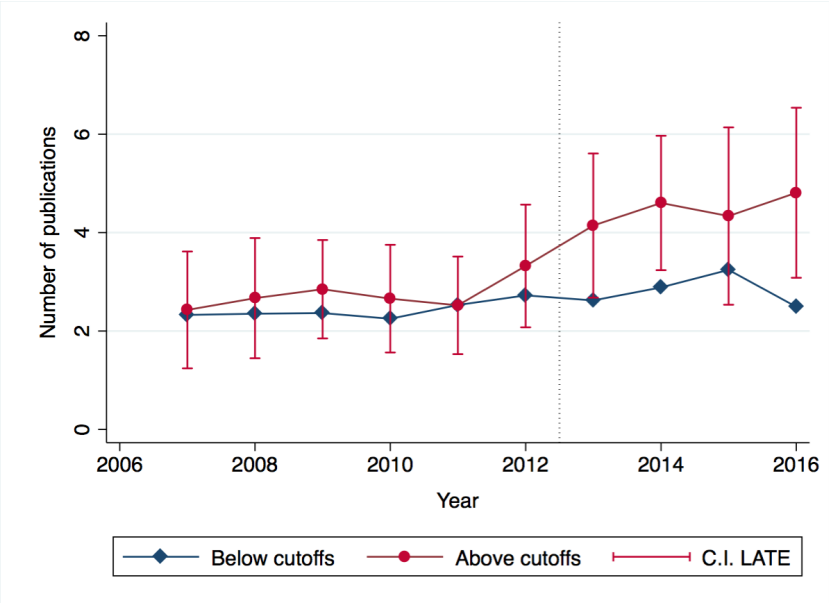
Figure 1.4: Discontinuity effect on the probability of success in the NSQ



a. N. of articles **b.** N. of citations **c.** H-index

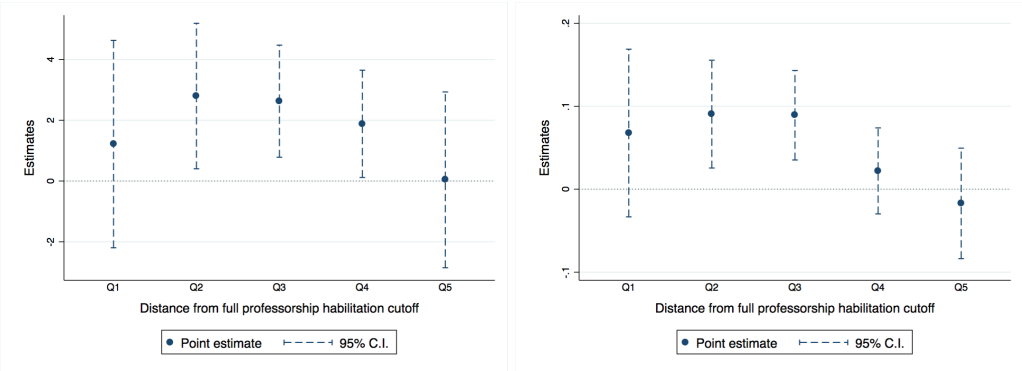
NOTES. This figure depicts the discontinuous jumps in the probability of achieving the qualification arising when each of the three indicators crosses the relevant cutoff value. The sample includes the candidates for an associate professorship in the 2012 NSQ in bibliometric fields. Each circle represents the average probability of achieving the NSQ within each 1-unit bin. The running variables for the three indicators are defined as the distance from the field-specific median, thus centered at zero. The dependent variable in the quadratic and local polynomial smoothing regression is an indicator that equals one when a candidate obtains the qualification. Both the quadratic and the local polynomial smoothing regressions are estimated within the interval of the relevant running variable [-20,20] in Panel A and B, while within the interval $[-10, 10]$ in Panel C. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations.

Figure 1.5: RD estimates of achieving the NSQ on the number of publications



NOTES. This figure depicts the OLS coefficients of the indicator for the compliance with the two-out-of-three rule on the quantity of publications estimated separately for each year over the 2007-2013 period. The sample includes the candidates for an associate professorship in the 2012 NSQ in bibliometric fields. Each blue dot corresponds to a weighted average of the field-specific dummies included in the regression for the specified year; each red dot corresponds to the LATE coefficient plus the 'weighted' constant term. Thus, the distance between the two dots represents the LATE for each regression. The LATE is the 2SLS coefficient of the indicator that equals one when a candidate achieves the qualification. The independent variable in all regressions is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. First-stage estimates are reported in Column (7) of Table 1.3. The dependent variables is the total number of papers (including articles, conference papers, reviews and other items) in the specified year. In each regression, within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.

Figure 1.6: Effect heterogeneity: distance from the full professor thresholds

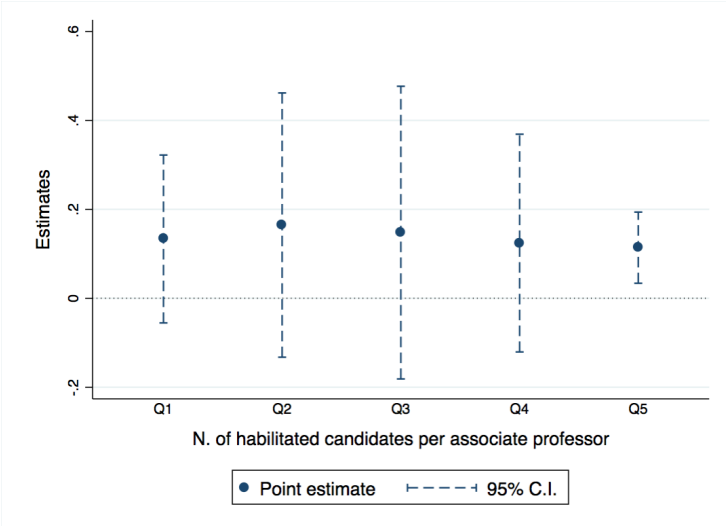


a. Number of publications (2013-2016)

b. Probability of achieving the full professor NSQ (2013-2016)

NOTES. This figure depicts the OLS coefficients of the interaction between the indicator for the compliance with the two-out-of-three rule and a set of indicators summarizing the heterogeneity in the distance from the full-professor thresholds on the quantity of publications over the 2013-2016 period. The sample includes the candidates for an associate professorship in the 2012 NSQ in bibliometric fields. The blue circles correspond to the coefficient of an interaction term formed by multiplying an indicator that equals one when a candidate complies with the two-out-of-three rule with a set of indicators that equal 1 if a candidate belongs to the specified quintile of the distribution of the distance from the full professor thresholds. The distance from the full professor threshold is expressed as a percentage of the initial stock of articles, thus defined as $dist_{i,1,s} = \frac{m_{1,s}^{full} - x_{i,1}}{x_{i,1}}$, where $m_{1,s}^{full}$ is the field-specific cutoff for the full professor NSQ and $x_{i,1}$ is the professor's score in the same indicator. The dependent variable is the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period. In each regression, within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.

Figure 1.7: Effect heterogeneity: degree of internal competition



NOTES. This figure depicts the OLS coefficients of the interaction between the indicator for compliance with the two-out-of-three rule and a set of indicators summarizing the heterogeneity in the degree of competition in the decentralized stage on the quantity of publications over the 2013-2016 period. The sample includes the candidates for an associate professorship in the 2012 NSQ in bibliometric fields. The blue circles correspond to the coefficient of an interaction term formed by multiplying an indicator that equals one when a candidate complies with the two-out-of-three rule with a set of indicators that equal one if a candidate belongs to the specified quintile of the distribution of the degree of competition in the decentralized stage. The degree of competition in the decentralized stage is defined as the ratio between the number of assistant professors with an associate professor NSQ and the number of associate professors in each field at the end of 2012. The dependent variable is the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period. In each regression, within each academic field we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and the ones with less than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.

Table 1.1: Descriptive Statistics

Panel A: applicants' characteristics		
	Mean	Sd
Academic age	15.87	7.65
Female	0.42	0.49
N. of applications	1.26	0.65
Qualified	0.58	0.49
Above 2/3 medians	0.66	0.47
Above median (n. of articles)	0.62	0.49
Above median (n. of citations)	0.68	0.47
Above median (h-index)	0.58	0.49
Distance from field median (n. of articles)	4.43	18.26
Distance from field median (n. of citations)	9.62	38.36
Distance from field median (h-index)	0.92	3.55
N. of scholars	4920	
Panel B: number of publications 2009-2011		
	Mean	Sd
Number of publications	12.49	13.76
Number of articles	9.34	11.70
Number of conference papers	1.70	4.26
Number of reviews	0.71	1.63
N. of scholars	4920	
Panel C: number of publications 2009-2011		
	Mean	Sd
Top 1% journal	0.02	0.15
CiteScore (percentile)	75.44	23.19
CiteScore	2.80	2.13
Sjr	1.52	1.58
Snip	1.31	0.91
Journal unlisted	0.13	0.34
N. of citations received	5.34	15.00
Single-authored	0.02	0.12
N. of authors	10.82	17.09
N. of publications	83670	

NOTES. This table reports the baseline characteristics of the candidates for an associate professorship in the 2012 NSQ in bibliometric fields. The unit of analysis is the single candidate Panels (A) and (B), while the single publication in (C). Panel (B) reports only scholars with at least one record in the *Scopus* database during the period 2013-2016. In all Panels, within each academic field we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations.

Table 1.2: Continuity test (2009-2011 measures)

Candidates for associate professor	(1) N. of publications	(2) Zero publications	(3) % Unlisted	(4) Top 5%	(5) CiteScore (pct)	(6) N. Coauthors (mean)	(7) Single-author	(8) Network size
ITT	0.392	-0.002	0.028	0.173	0.010	0.310	-0.013	1.372
	(0.467)	(0.007)	(0.034)	(1.143)	(0.010)	(0.315)	(0.025)	(3.608)
Mean Dep. Var.	11.658	0.026	0.580	70.999	0.122	6.860	0.143	41.859
Standard dev.	9.598	0.159	0.494	15.860	0.196	5.766	0.351	53.641
N. of clusters	82	82	82	82	82	82	82	82
Observations	4920	4920	4755	4755	4792	4763	4861	4785

NOTES. This table reports the Intention-to-treat (ITT) and the Local Average Treatment Effect (LATE) of achieving the qualification on the quantity of publications, their quality and the number of collaborations computed over the period 2009-2011. The sample includes the candidates for an associate professorship in the 2012 NSQ in bibliometric fields. The dependent variables are: the total number of papers – including articles, conference papers, reviews and other items– (Column 1); an indicator that equals one when a candidate does not publish any paper (Column 2); the share of publications in journals not classified in the *Scopus* database (Column 3); the share of articles published in journals scoring in the top 5% according to the 2012 *CiteScore journal percentile* (Column 4); the average *CiteScore journal percentile* (Column 5); the average number of co-authors per publication (Column 6); the share of single-authored publications (Column 7) and the total number of distinct co-authors (Column 8). In all columns the dependent variable is referred to the period 2009-2011. The ITT is the OLS coefficient of the indicator that equals one when a candidate complies with the two-out-of-three rule; the LATE is the 2SLS coefficient of the indicator that equals one when a candidate gets the qualification. The independent variable in this regression is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. First stage estimates are reported in Columns (5) and (6) of Table 1.3. In Column (6) the sample includes only scholars with at least one record in the *Scopus* database during the period 2013-2016. In Columns (3) to (5) the sample is further limited to scholars with at least one publication in a journal classified in the *Scopus* database (with a non-missing score) during the period 2009-2011. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level, in parentheses. *** $p< 0.01$, ** $p<0.05$, and *$p<0.10$.

Table 1.3: First stage

Candidates for Associate professor							
	Single RD (Articles)		Single RD (Citations)		Single RD (H-Index)		Triple RD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Quadratic	LLR	Quadratic	LLR	Quadratic	LLR	Quadratic
Above cutoff	0.131***	0.122***	0.210***	0.161***	0.153***	0.162**	
	(0.040)	(0.046)	(0.041)	(0.057)	(0.046)	(0.066)	
Above 2/3 cutoffs							0.306***
							(0.043)
Academic field FE	Yes	No	Yes	No	Yes	No	Yes
Field specific interactions	Yes	No	Yes	No	Yes	No	Yes
Mean dep. var.	0.577	0.573	0.577	0.536	0.577	0.564	0.577
Sd dep. var.	0.494	0.495	0.494	0.499	0.494	0.496	0.494
BW (MSE)		0.396		0.364		0.304	
N of clusters	82	89	82	89	82	89	82
Observations	4920	2752	4920	1798	4920	3034	4920

NOTES. This table reports the OLS coefficients of overcoming the field-specific bibliometric cutoffs on the outcome of the 2012 NSQ. The sample includes the candidates for an associate professorship in the 2012 NSQ in bibliometric fields. In all columns, the dependent variable is an indicator that equals one when a candidate obtains the qualification, and zero otherwise. In Columns (1) to (6), the main independent variable is an indicator that equals one when a candidate overcomes the relevant cutoff for either the number of articles (Columns 1 and 2), the number of citations (Columns 3 and 4) and the h-index (Columns 5 and 6). In Column (7), the main independent variable is an indicator that equals one when a candidate complies with the two-out-of-three rule, that is, when her scores in at least two indicators are above the relevant cutoffs. Regressions in Columns (1), (3), (5) and (6) are estimated using a field-specific polynomial (quadratic) specification over the entire support after excluding observations in the top 10% and the bottom 1% of the within-field distribution of the distances from each cutoff. In Columns (2), (4) and (6), we replicate the estimates in (1), (3), and (5), performing local linear regressions (LLR) within the MSE-optimal bandwidths computed following Calonico et al. (2014) – using the companion Stata package described in Calonico et al. (2017) – after normalizing each distance from the cutoff by dividing it by the cutoff itself. In all specifications, we exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. Standard errors, clustered at the academic field level, in parentheses. *** $p< 0.01$, ** $p<0.05$, and *$p<0.10$.

Table 1.4: The effect of achieving the NSQ on the number of published papers

Candidates for Associate professor				
	Publications	Articles	Conf. Papers	Reviews
	(1)	(2)	(3)	(4)
ITT	2.003***	1.254**	0.540	0.198*
	(0.701)	(0.477)	(0.336)	(0.102)
LATE	6.557***	4.105***	1.769*	0.648**
	(2.159)	(1.485)	(0.956)	(0.290)
Mean Dep. Var.	17.006	12.747	2.313	0.916
Standard dev.	20.410	17.820	5.719	2.088
N. of clusters	82	82	82	82
Observations	4920	4920	4920	4920

NOTES. This table reports the ITT and the LATE of achieving the qualification on the quantity of publications over the 2013-2016 period. The sample includes the candidates for an associate professorship in the 2012 NSQ in bibliometric fields. The dependent variable in Column (1) is the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period; the dependent variables in Columns (2), (3) and (4) are the total number of articles, conference papers and reviews published during the 2013-2016 period, respectively. The ITT is the OLS coefficient of the indicator that equals one when a candidate complies with the two-out-of-three rule; the LATE is the 2SLS coefficient of the indicator that equals one when a candidate obtains the qualification. The independent variable in this regression is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. First-stage estimates are reported in Column (7) of Table 1.3. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level, in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$.

Table 1.5: Effect heterogeneity: gender

Dependent variable: Number of publications		
	First stage	LATE
	(1)	(2)
female	0.005	-1.148
	(0.020)	(0.875)
above 2/3 cutoffs	0.329***	
	(0.046)	
above 2/3 cutoffs × female	-0.057*	
	(0.030)	
qualified		6.742***
		(2.253)
qualified × female		-0.684
		(1.563)
Mean Dep. Var.	0.577	17.006
Standard dev.	0.494	20.410
N. of clusters	82	82
Observations	4920	4920

NOTES. This table reports the OLS coefficient of the indicator for the compliance with the two-out-of-three rule and its interaction with gender on the outcome of the 2012 NSQ (Column 1) and on the quantity of publications over the period 2013-2016 (Column 2). The sample includes the candidates for an associate professorship in the 2012 NSQ in bibliometric fields. The dependent variable in Column (1) is an indicator that equals one when a candidate obtains the qualification, and zero otherwise; the dependent variable in Column (2) is the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period. In both columns, the main independent variables are an indicator that equals one when a candidate complies with the two-out-of-three rule, an indicator that equals one in case of a female candidate and an interaction term formed by multiplying the two indicators. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level, in parentheses. *** $p < 0.01$, ** $p < 0.05$, and *$p < 0.10$.

Table 1.6: The effect of achieving the NSQ on the number of citations received

Candidates for Associate professor				
	Total citations	Mean citations	% Cit. $\geq$ 50	% Zero cit.
ITT	13.526**	0.519*	0.002	0.012
	(6.372)	(0.273)	(0.002)	(0.012)
LATE	43.934**	1.687**	0.006	0.039
	(18.087)	(0.748)	(0.004)	(0.034)
Mean Dep. Var.	92.371	4.417	0.005	0.356
Standard dev.	200.847	4.184	0.028	0.211
N. of clusters	82	82	82	82
Observations	4838	4838	4838	4838

NOTES. This table reports the ITT and the LATE of achieving the qualification on the citations received by papers published over the 2013-2016 period. The sample includes the candidates for an associate professorship in the 2012 NSQ in bibliometric fields with at least one record in the *Scopus* database over the 2013-2016 period. The dependent variables in Columns (1) and (2) are the total and the average number of citations received by papers published during the 2013-2016 period, respectively. The dependent variable in Columns (3) and (4) are the share of papers published during the 2013-2016 period with at least 50 citations and with zero citations, respectively. The ITT is the OLS coefficient of the indicator that equals one when a candidate complies with the two-out-of-three rule; the LATE is the 2SLS coefficient of the indicator that equals one when a candidate obtains the qualification. The independent variable in this regression is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. First-stage estimates are reported in Column (7) of Table 1.3. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level, in parentheses. *** p$<$ 0.01, ** p$<$0.05, and *p$<$0.10.

Table 1.7: The effect of achieving the NSQ on the quality of publications

Candidates for Associate professor						
	Top 5%	CiteScore (pct)	CiteScore	Sjr	Snip	% Unlisted
	(1)	(2)	(3)	(4)	(5)	(6)
ITT	0.038	-0.110	0.063	0.052	0.040	0.000
	(0.030)	(1.087)	(0.069)	(0.059)	(0.032)	(0.011)
LATE	0.126	-0.360	0.206	0.170	0.130	0.001
	(0.083)	(3.065)	(0.192)	(0.168)	(0.086)	(0.029)
Mean Dep. Var.	0.655	72.403	2.509	1.336	1.239	0.126
Standard dev.	0.475	14.773	1.343	0.843	0.456	0.180
N. of clusters	82	82	82	82	82	82
Observations	4809	4809	4809	4808	4809	4838

NOTES. This table reports the ITT and the LATE of achieving the qualification on the quality of publications over the 2013-2016 period. The sample includes the candidates for an associate professorship in the 2012 NSQ in bibliometric fields. The dependent variable in Column (1) is the share of articles published in journals scoring in the top 5% according to the 2015 *CiteScore journal percentile* during the 2013-2016 period; the dependent variables in Columns (2), (3), (4) and (5) are the average *CiteScore journal percentile*, the average *CiteScore* index, the average *Sjr* index, the average *Snip* index of papers published during the 2013-2016 period, respectively; the dependent variable In Column (6) is the share of publications in journals not classified in the *Scopus* database during the period 2013-2016. The ITT is the OLS coefficient of the indicator that equals one when a candidate complies with the two-out-of-three rule; the LATE is the 2SLS coefficient of the indicator that equals one when a candidate obtains the qualification. The independent variable in this regression is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. First-stage estimates are reported in Column (7) of Table 1.3. In Column (6) the sample includes only scholars with at least one record in the *Scopus* database during the period 2013-2016. In Columns (2) to (5) the sample is further limited to scholars with at least one publication in a journal classified in the *Scopus* database of journals (with a non-missing score) during the period 2013-2016. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and the ones with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level, in parentheses. *** $p< 0.01$, ** $p<0.05$, and *$p<0.10$.

Table 1.8: The effect of achieving the NSQ on the number of collaborations

Candidates for Associate professor				
	(1)	(2)	(3)	(4)
	N. Coauthors	N. Coauthors	Single-	Network
	(mean)	(median)	-author	size
ITT	0.241	0.686	0.016	1.748
	(0.459)	(0.419)	(0.029)	(5.334)
LATE	0.785	2.236*	0.053	5.722
	(1.309)	(1.292)	(0.082)	(15.061)
Mean Dep. Var.	8.185	7.255	0.137	62.016
Standard dev.	8.245	8.217	0.344	84.882
N. of clusters	82	82	82	82
Observations	4801	4801	4801	4806

NOTES. This table reports the ITT and the LATE of achieving the qualification on the number of co-authorships over the 2013-2016 period. The sample includes the candidates for an associate professorship in the 2012 NSQ in bibliometric fields with at least one record in the *Scopus* database over the 2013-2016 period. The dependent variables are the average number of coauthors per publication (Column 1), the maximum number of co-authors per publication (Column 2), the share of single-authored publications (Column 3) and the total number of distinct coauthors (Column 4). The ITT is the OLS coefficient of the indicator that equals one when a candidate complies with the two-out-of-three rule; the LATE is the 2SLS coefficient of the indicator that equals one when a candidate obtains the qualification. The independent variable in this regression is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. First-stage estimates are reported in Column (7) of Table 1.3. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and the ones with less than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level, in parentheses. *** $p< 0.01$, ** $p<0.05$, and *$p<0.10$.

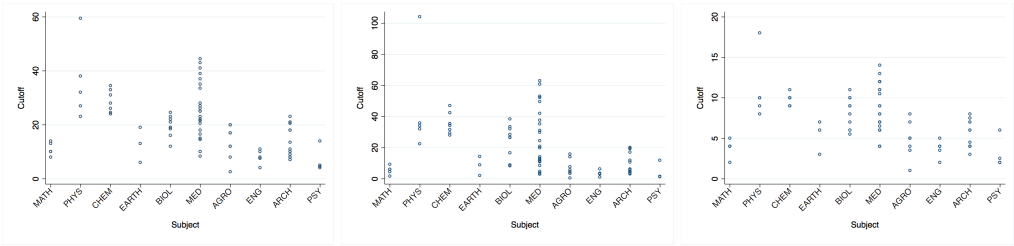
Table 1.9: The effect of achieving the full professor NSQ on the number of published papers

Candidates for Full professor				
	Publications	Articles	Conf. Papers	Reviews
	(1)	(2)	(3)	(4)
ITT	-0.349	-0.251	0.195	-0.105
	(1.195)	(1.009)	(0.221)	(0.189)
LATE	-0.836	-0.601	0.468	-0.251
	(2.408)	(2.037)	(0.467)	(0.382)
Mean Dep. Var.	22.592	16.776	2.727	1.431
Standard dev.	24.987	21.383	6.982	2.758
N. of clusters	51	51	51	51
Observations	2746	2746	2746	2746

NOTES. This table reports the ITT and the LATE of achieving the qualification on the quantity of publications over the 2013-2016 period. The sample includes the candidates for a full professorship in the 2012 NSQ in bibliometric fields. The dependent variable in Column (1) is the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period; the dependent variables in Columns (2), (3) and (4) are the total number of articles, conference papers and reviews published during the 2013-2016 period, respectively. The ITT is the OLS coefficient of the indicator that equals one when a candidate complies with the two-out-of-three rule; the LATE is the 2SLS coefficient of the indicator that equals one when a candidate obtains the qualification. The independent variable in this regression is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. First-stage estimates are reported in Column (7) of Table 1.3. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level, in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$.

1.10 Appendix

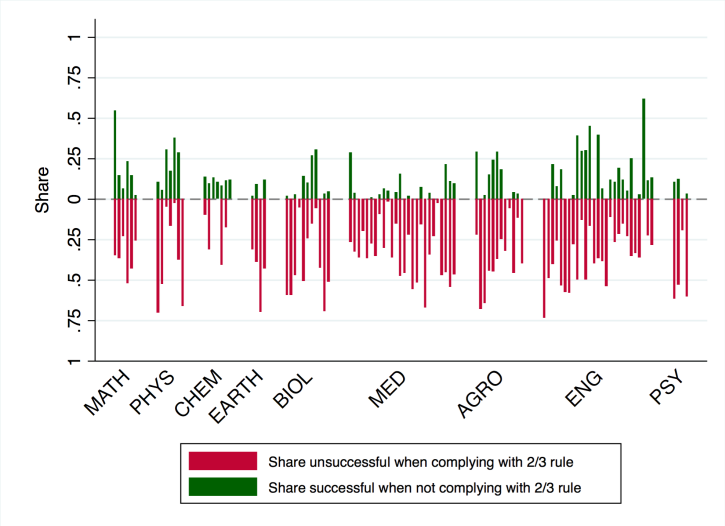
Figure 1.8: Heterogeneity in field-specific cutoffs



a. Number of articles **b.** Number of citations **c.** H-index

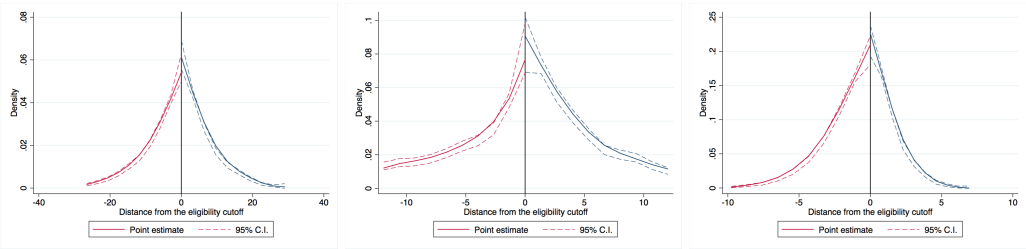
NOTES. This figure depicts the value of the cutoffs for each bibliometric indicator and academic field. Each circle corresponds to an academic field, grouped according to its main subject. We exclude the fields with more than 90% successful candidates and those with fewer than 30 observations.
Legend: MATH=Mathematics; PHYS=Physics; CHEM=Chemistry; EARTH=Earth Sciences; BIOL= Biology; MED=Health Sciences; AGRO=Agronomy and Veterinary; ENG=Engineering; ARCH=Architecture; PSY=Psychology.

Figure 1.9: Compliance with the two-out-of-three rule



NOTES. This figure depicts the share of candidates for an associate professorship in the 2012 NSQ in bibliometric fields, depending on the compliance with the two-out-of-three rule and the outcome of the 2012 NSQ. Each bar corresponds to an academic field, grouped according to its main subject. The length of each red (green) bar indicates the share of candidates who did not obtain (obtained) the qualification in the 2012 NSQ, even when complying (not complying) with the two-out-of-three rule. We exclude the fields with more than 90% successful candidates and those with fewer than 30 observations.
Legend: MATH=Mathematics; PHYS=Physics; CHEM=Chemistry; EARTH=Earth Sciences; BIOL= Biology; MED=Health Sciences; AGRO=Agronomy and Veterinary; ENG=Engineering; ARCH=Architecture; PSY=Psychology.

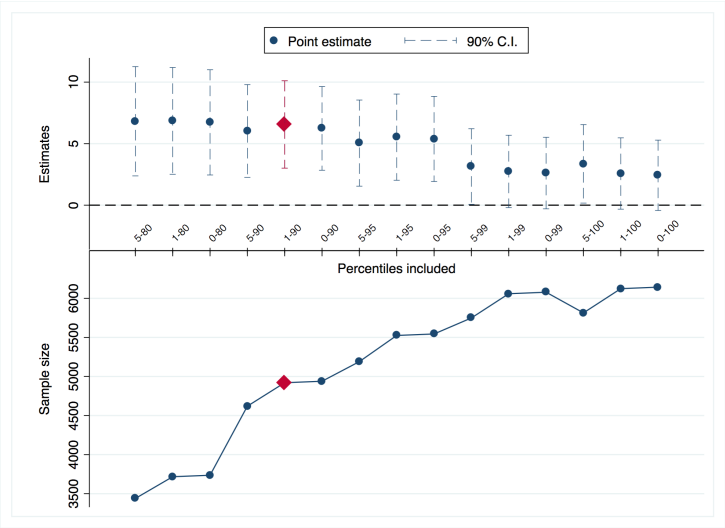
Figure 1.10: Manipulation testing



RV: n. of articles RV: n. of citations RV: H-index

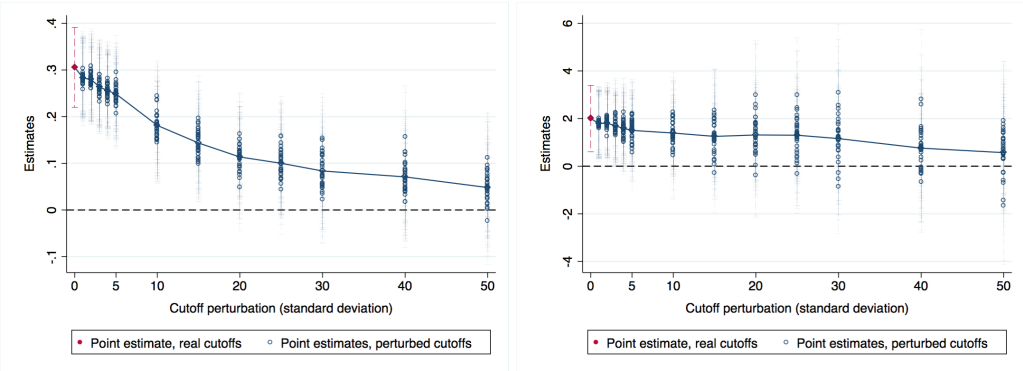
NOTES. This figure depicts the local polynomial density estimation of candidates for an associate professorship in the 2012 NSQ in bibliometric fields, depending on their distance from each of the three bibliometric cutoffs. The local polynomial density is estimated following Cattaneo et al. (2017b) and using the companion Stata package described in Cattaneo et al. (2017a).

Figure 1.11: Robustness to sample restrictions



NOTES. This figure depicts, in the upper panel, the OLS coefficients of the indicator for the compliance with the two-out-of-three rule on the quantity of publications over the 2013-2016 period under different sample restrictions. The sample, before imposing the different restrictions, includes the candidates for an associate professorship in the 2012 NSQ in bibliometric fields. Each dot in the upper panel corresponds to the ITT coefficients estimated on specified inter-percentile ranges computed for candidates' distance from the relevant cutoffs. In each regression, the dependent variable is the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period. The main independent variable is an indicator that equals one when a candidate complies with the two-out-of-three rule, that is, when her scores in at least two indicators are above the relevant cutoffs. The different sample restrictions are reported on the x-axis. For instance, when estimating the regression within the inter-percentile range '20-90', we exclude all candidates belonging to the top 10% or the bottom 20% of the pool of applicants in the same field for any of the three indicators considered. In the lower panel, the figures reports the sample size under the different sample restrictions. The red dots correspond to the sample chosen in our baseline specification. In each regression, within each academic field we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.

Figure 1.12: Placebo test: cutoff perturbation

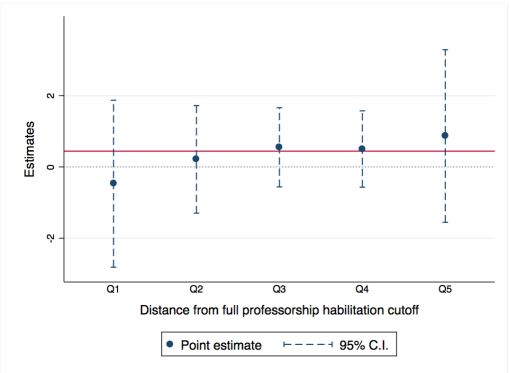


a. First stage

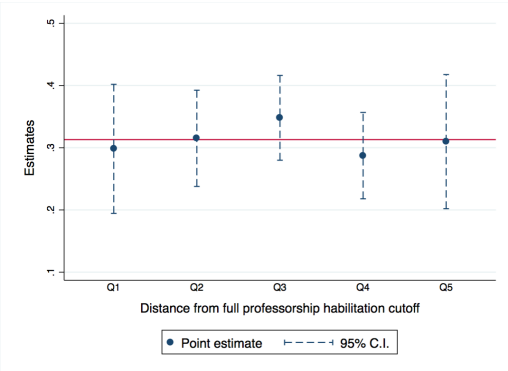
b. Number of publications 2013-2016 (ITT)

NOTES. This figure depicts the OLS coefficients of the indicator for the compliance with the two-out-of-three rule on the outcome of the 2012 NSQ (Panel a) and on the quantity of publications over the 2013-2016 period (Panel B) applying different perturbations to the cutoff values. The sample includes the candidates for an associate professorship in the 2012 NSQ in bibliometric fields. Each dot corresponds to the coefficient of an indicator that equals one when a candidate complies with the two-out-of-three rule on the depended variable considered, under a different perturbation of the cutoff values. The dependent variables in Panels A and B are an indicator that equals one when a candidate obtains the qualification and the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period, respectively. The permutations of the cutoff values are obtained by adding a randomly generated error component $\epsilon \sim N(0, \sigma)$, where the standard deviation (σ) of the error determines the intensity of the reshuffling. For each value of σ, we estimate 30 separate regressions for different realizations of ϵ. We apply the same reshuffling to the three bibliometric cutoffs, and we force the perturbation to lie within within - and +100% of the original cutoff values. In each regression, within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.

Figure 1.13: Effect heterogeneity: distance from the full-professor thresholds



a. Number of publications (2009-2012)



b. Probability of achieving the associate professor NSQ (2012)

NOTES. This figure depicts the OLS coefficients of the interaction between the indicator for the compliance with the two-out-of-three rule and a set of indicators summarizing the heterogeneity in the distance from the full-professor thresholds on the quantity of publications over the 2013-2016 period. The sample includes the candidates for an associate professorship in the 2012 NSQ in bibliometric fields. The blue circles correspond to the coefficient of an interaction term formed by multiplying an indicator that equals one when a candidate complies with the two-out-of-three rule with a set of indicators that equal one if a candidate belongs to the specified quintile of the distribution of the distance from the full professor thresholds. The distance from the full professor threshold is expressed as a percentage of the initial stock of articles, thus defined as $dist_{i,1,s} = \frac{m_{1,s}^{full} - x_{i,1}}{x_{i,1}}$, where $m_{1,s}^{full}$ is the field-specific cutoff for the full professor NSQ and $x_{i,1}$ the professor's score in the same indicator. The dependent variable is the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period. In each regression, within each academic field we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.

Table 1.10: First stage - full professor NSQ

Candidates for full professor							
	Single RD (Articles)		Single RD (Citations)		Single RD (H-Index)		Triple RD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Quadratic	LLR	Quadratic	LLR	Quadratic	LLR	Quadratic
Above cutoff	0.141*	0.175***	0.242***	0.174**	0.265***	0.312***	
	(0.072)	(0.064)	(0.075)	(0.081)	(0.086)	(0.108)	
Above 2/3 cutoffs							0.420***
							(0.067)
Academic field FE	Yes	No	Yes	No	Yes	No	Yes
Field specific interactions	Yes	No	Yes	No	Yes	No	Yes
Mean dep. var.	0.566	0.581	0.566	0.547	0.566	0.543	0.566
Sd dep. var.	0.496	0.494	0.496	0.498	0.496	0.498	0.496
BW (MSE)		0.298		0.270		0.182	
N of clusters	47	89	47	89	47	89	47
Observations	2369	1039	2369	675	2369	905	2369

NOTES. This table reports the OLS coefficients of overcoming the field-specific bibliometric cutoffs on the outcome of the 2012 NSQ. The sample includes the candidates for a full professorship in the 2012 NSQ in bibliometric fields. In all columns, the dependent variable is an indicator that equals one when a candidate gets the qualification, and zero otherwise. In Columns (1) to (4) the main independent variable is an indicator that equals one when a candidate overcomes the relevant cutoff for either the number of articles (Columns 1 and 2) or the number of citations (Columns 3 and 4). As for the h-index, estimates are not reported in this table. In Columns (5) and (6) the main independent variable is an indicator that equals one when a candidate complies with the two-out-of-three rule, that is, when her scores in at least two indicators are above the relevant cutoffs. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% of successful candidatesand those with fewer than 30 observations. Regressions in Columns (1), (3), (5) and (6) are estimated using a polynomial (quadratic) specification over the entire support; in Columns (2) and (4) Local Linear Regressions (LLR) are estimated within the MSE-optimal bandwidth computed following Calonico et al. (2014), using the companion Stata package described in Calonico et al. (2017). In Column (6) Local Linear Regression is estimated within an arbitrary bandwidth, as the MSE-optimal bandwidth cannot be computed.
Standard errors, clustered at the academic field level, in parentheses. *** $p< 0.01$, ** $p<0.05$, and *$p<0.10$.

Table 1.11: Robustness check: local linear specification

Dependent variable: Number of publications		
	(1)	(2)
	Full Parametric	LLR
ITT	2.003***	1.559*
	(0.701)	(0.815)
LATE	6.557***	6.620*
	(2.159)	(3.718)
First Stage	0.306***	0.236***
	(0.043)	(0.056)
Academic field FE	Yes	Yes
Field specific interactions	Yes	No
Mean Dep. Var.	17.006	15.392
Standard dev.	20.410	11.950
BW (MSE) Ind 1		0.349
BW (MSE) Ind 2		0.465
BW (MSE) Ind 3		0.230
N. of clusters	82	84
Observations	4920	931

NOTES. This table reports the ITT and the LATE of achieving the qualification on the quantity of publications over the 2013-2016 period, and the corresponding first-stage estimates. In Column (1) we use our baseline full-parametric specification over the entire support. In Column (2), we use a linear approach in the neighborhood of the threshold. The sample includes the candidates for an associate professorship in the 2012 NSQ in bibliometric fields. The dependent variable in both columns is the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period. The ITT is the OLS coefficient of the indicator that equals one when a candidate complies with the two-out-of-three rule; the LATE is the 2SLS coefficient of the indicator that equals one when a candidate obtains the qualification. The independent variable in this regression is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. The polynomial (quadratic) specification in Column (1) is estimated over the entire support after excluding observations in the top 10% and the bottom 1% of the distribution of the distances. The local linear specification in Column (2) is estimated within a joint three-dimensional bandwidth. The bandwidth for each productivity indicator is the MSE-optimal bandwidth computed following Calonico et al. (2014) and using the companion Stata package described in Calonico et al. (2017). In both columns, we exclude the fields with more than 90% successful candidates and those with fewer than 30 observations.
Standard errors, clustered at the academic field level, in parentheses. *** $p< 0.01$, ** $p<0.05$, and *$p<0.10$.

Table 1.12: Robustness check: candidates applying to the competition sector they belong to as of 2012.

Candidates for associate professor				
	Publications	Articles	Conf. Papers	Reviews
	(1)	(2)	(3)	(4)
ITT	1.520**	1.230**	0.218	0.045
	(0.761)	(0.533)	(0.290)	(0.109)
LATE	4.238**	3.431**	0.609	0.127
	(1.951)	(1.376)	(0.709)	(0.263)
Mean Dep. Var.	15.374	11.285	2.216	0.895
Standard dev.	13.728	10.111	5.719	2.048
N. of clusters	83	83	83	83
Observations	5024	5024	5024	5024

NOTES. This table reports the Intention-to-treat (ITT) and the Local Average Treatment Effect (LATE) of achieving the qualification on the quantity of publications over the period 2013-2016. The sample includes the candidates for an associate professorship in the 2012 NSQ in bibliometric fields. In case of multiple applications, we consider the one to the competition sector to which the scholar already belongs as an assistant professor as of December 2012. The dependent variable in Column (1) is the total number of papers (including articles, conference papers, reviews and other items) published during the period 2013-2016; the dependent variables in Columns (2), (3) and (4) are the total number of articles, conference papers and reviews published during the period 2013-2016, respectively. The ITT is the OLS coefficient of the indicator that equals one when a candidate complies with the two-out-of-three rule; the LATE is the 2SLS coefficient of the indicator that equals one when a candidate gets the qualification. The independent variable in this regression is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level, in parentheses. *** $p< 0.01$, ** $p<0.05$, and *$p<0.10$.

Table 1.13: The effect of achieving the NSQ on the number of publications, per year

Panel A - Dependent variable: Number of publications								
	2009	2010	2011	2012	2013	2014	2015	2016
ITT	0.149	0.138	0.008	0.253	0.474**	0.496**	0.323	0.717***
	(0.154)	(0.167)	(0.153)	(0.222)	(0.236)	(0.211)	(0.273)	(0.254)
LATE	0.486	0.452	0.025	0.827	1.556**	1.626***	1.058	2.348***
	(0.446)	(0.482)	(0.430)	(0.634)	(0.679)	(0.618)	(0.801)	(0.783)
Mean Dep. Var.	2.940	2.962	3.146	3.533	3.820	4.159	4.518	4.574
Standard dev.	3.015	3.019	3.141	3.793	3.862	4.956	7.547	8.446
Observations	4827	4843	4855	4861	4869	4898	4914	4919
Panel B - Dependent variable: No publications								
	2009	2010	2011	2012	2013	2014	2015	2016
ITT	-0.043	-0.029	0.011	-0.029	-0.050	-0.019	-0.005	-0.032
	(0.027)	(0.031)	(0.029)	(0.022)	(0.030)	(0.023)	(0.028)	(0.026)
LATE	-0.140*	-0.094	0.036	-0.096	-0.162*	-0.063	-0.015	-0.103
	(0.078)	(0.092)	(0.079)	(0.063)	(0.086)	(0.066)	(0.079)	(0.074)
Mean Dep. Var.	0.147	0.160	0.153	0.125	0.136	0.123	0.128	0.129
Standard dev.	0.354	0.366	0.360	0.331	0.343	0.328	0.334	0.336
Observations	4920	4920	4920	4920	4920	4920	4920	4920
Panel C - Dependent variable: Number of articles								
	2009	2010	2011	2012	2013	2014	2015	2016
ITT	0.056	0.097	0.079	0.169	0.231	0.342*	0.160	0.530***
	(0.120)	(0.117)	(0.114)	(0.165)	(0.161)	(0.174)	(0.204)	(0.186)
LATE	0.183	0.316	0.258	0.555	0.758	1.121**	0.526	1.735***
	(0.338)	(0.336)	(0.326)	(0.474)	(0.468)	(0.514)	(0.590)	(0.567)
Mean Dep. Var.	2.084	2.041	2.206	2.468	2.734	3.054	3.408	3.598
Standard dev.	2.339	2.229	2.327	2.979	2.849	4.183	6.854	7.807
Observations	4827	4843	4855	4861	4869	4898	4914	4919

NOTES. This table reports the ITT and the LATE of achieving the NSQ on the quantity of publications in each year over the 2009-2016 period. The sample includes the candidates for an associate professorship in the 2012 NSQ in bibliometric fields. The dependent variables in Panels A and C are the numbers of publications and articles, respectively. The dependent variable in Panel B is an indicator that equals one when a candidate does not publish any paper in a given year. The ITT is the OLS coefficient of the indicator that equals one when a candidate complies with the two-out-of-three rule; the LATE is the 2SLS coefficient of the indicator that equals one when a candidate achieves the NSQ. The independent variable in this regression is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. First-stage estimates are reported in Column (7) of Table 1.3. In Panels A and C, the sample includes only scholars that were 'active' in the year considered, that is, scholars whose first publication is not later than that year. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level (N=82), in parentheses. *** $p< 0.01$, ** $p<0.05$, and *$p<0.10$.

Table 1.14: Effect heterogeneity: main subject

Dependent variable: Number of publications										
	MATH	PHYS	CHEM	EARTH	BIOL	MED	AGRO	ENG	ARCH	PSY
Panel A: LATE										
qualified × subject	1.338	50.127***	13.437*	8.069**	0.082	3.582	19.019	20.410*	19.064	-3.047
	(2.764)	(8.913)	(7.451)	(3.865)	(1.317)	(2.476)	(18.982)	(11.371)	(30.293)	(4.808)
Mean (subject)	11.646	39.790	17.721	13.601	11.638	16.691	12.519	14.379	21.150	11.897
St. dev. (subject)	11.014	56.737	10.948	7.949	8.194	15.777	9.470	10.079	15.106	10.291
Panel B: First Stage										
above 2/3 cutoffs × subject	0.372***	0.143***	0.242***	0.507**	0.313***	0.394***	0.186*	0.252	0.114	0.532**
	(0.109)	(0.051)	(0.064)	(0.243)	(0.074)	(0.096)	(0.105)	(0.206)	(0.140)	(0.215)
Mean (subject)	0.553	0.790	0.703	0.710	0.523	0.445	0.672	0.551	0.703	0.623
St. dev. (subject)	0.498	0.408	0.458	0.455	0.500	0.497	0.470	0.499	0.458	0.486

NOTES. This table reports the OLS coefficients of the indicator for compliance with the two-out-of-three rule and its interaction with a set of subject-specific dummies on the quantity of publications over the 2013-2016 period (Panel A) and on the outcome of the 2012 NSQ (Panel B). The sample includes the candidates for an associate professorship in the 2012 NSQ in bibliometric fields. The dependent variable in Panel A is the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period; the dependent variable in Panel B is an indicator that equals one when a candidate obtains the qualification, and zero otherwise. In both panels, the main independent variables are a set of interaction terms formed by multiplying an indicator that equals one when a candidate complies with the two-out-of-three rule with each of the subject-specific dummies, that is, a set of indicators equal to 1 if a candidate belongs to the specified subject. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level, in parentheses. *** $p< 0.01$, ** $p<0.05$, and *$p<0.10$.
Legend: MATH=Mathematics; PHYS=Physics; CHEM=Chemistry; EARTH=Earth Sciences; BIOL= Biology; MED=Health Sciences; AGRO=Agronomy and Veterinary; ENG=Engineering; ARCH=Architecture; PSY=Psychology.

Chapter 2

CUTTING THROUGH THE FOG: FINANCIAL LITERACY AND THE SUBJECTIVE VALUE OF FINANCIAL ASSETS

2.1 Introduction

In recent years, a growing emphasis has been placed on the role of financial literacy in individual decision-making.[1] Financial literacy potentially affects financial behavior in many ways. Saving and borrowing decisions, mortgage choices, stock market participation, and retirement planning are only some of the dimensions on which the agents' degree of financial sophistication can have a large impact. In spite of its importance for individual well-being, however, financial literacy is found to be alarmingly low even in developed countries, where most people have

[1]We refer to financial literacy as the level of understanding of basic financial concepts. Alternative and more complete definitions are given by Lusardi and Mitchell (2014) according to whom financial literacy is "peoples' ability to process economic information and make informed decisions about financial planning, wealth accumulation, debt, and pensions." (p.6). Similarly, Atkinson and Messy (2012) define it as "[...] a combination of awareness, knowledge, skill, attitude and behavior necessary to make sound financial decisions and ultimately achieve individual financial wellbeing." (p.14).

limited understanding of financial markets.[2]

Given the importance of the topic from a policy and academic perspective, numerous researchers analyze the relation between financial literacy and several individual financial outcomes. The resulting empirical evidence, although abundant, is mixed. While some authors find financial literacy to be effectively correlated with different financial habits, others fail to detect systematic effects of financial education programs in improving agents' financial decisions. The consequent lack of consensus about the impact of financial literacy calls for additional evidence and especially, for a deeper insight into the channels through which financial literacy can affect agents' financial behavior. This is precisely the aim of our analysis, which focuses on a so-far unexplored dimension of individual decision-making that can potentially be influenced by an agent's degree of financial sophistication: the subjective valuation of financial assets. Investigating whether financial literacy affects the way individuals perceive and evaluate financial assets can indeed improve our understanding of its impact on the financial behavior of households.

From a methodological point of view, assessing the causal impact of financial literacy is challenging. The main difficulty in this context is the potential endogeneity of financial literacy to the financial behavior of individuals. For instance, it is hard to disentangle whether financial market participation is the result of a higher level of financial sophistication or, *vice versa*, whether people who invest more in financial markets end up being more financially literate. In this study, we overcome this endogeneity problem by designing a randomized experiment – in a sample of 260 young adults in Spain – in which we randomly induce an exogenous increase in the level of financial literacy of a subsample of the participants before asking them to evaluate a risky lottery. Specifically, participants are exposed to a double randomized treatment, thus being split into four groups. In a setup *à la* Holt and Laury (2002), respondents have to make twenty subsequent choices between the risky lottery and an increasing safe amount of money. We measure the value each respondent assigns to the lottery – that is, its certainty equivalent – as the safe amount for which she stops accepting the gamble and switches to the safe alternative. The risky lottery is presented either as a simple coin flip or as a financial asset, in this case being framed with financial concepts. Importantly, the payoffs and the probabilities are exactly the same for all participants, regardless of the way the lottery is presented. Then, we induce an exogenous increase in the level of financial literacy of half of the participants exposed to each of the two framings by explaining basic financial notions to them. Providing such teaching treatment to participants in the coin-toss group as well allows us to disentangle the

[2]Evidence on households' illiteracy was firstly provided by Bernheim (1995) with US data. More recently, Atkinson and Messy (2012) and Klapper et al. (2015) present analogous findings with data from the OECD INFE Pilot Study and the Standard & Poor's Global Financial Literacy Survey, respectively.

effect of increasing financial literacy from any other behavioral response that the *teaching* might induce. In this setting, we measure the impact of both the financial framing and financial literacy on individuals' choices by comparing the average certainty equivalent in the four groups.

Two main results emerge from the experiment: i) the financial framing of the lottery makes it less desirable than an equivalent coin toss to individuals who do not receive the *teaching*; ii) the increase in financial literacy provided through the *teaching* has a positive and significant impact on the average value assigned to the financially framed lottery. Importantly, no effect of the *teaching* emerges among participants evaluating the coin toss, thus confirming that receiving the *teaching* does not alter individual behavior when the choice does not involve any financial concept. We also find that both the financial framing and the *teaching* affect the average comprehension of the structure of the lottery. Respondents report a lower understanding of the risky option when this is presented as a financial asset and are less able to correctly compute its maximum and expected gain. Enhancing financial literacy – through the provision of the *teaching* treatment – effectively improves both the self-assessed and the actual understanding of the financial lottery.

This evidence documents that a lack of financial knowledge leads to a systematic undervaluation of financial products that would otherwise be desirable to investors. Promoting financial literacy might reduce this distortion by improving the understanding of financial assets, thus increasing the value households assign to them. When put together, our experimental results have important policy implications. They show that financial literacy and an adequate comprehension of financial products play a major role at the time of investment decisions. Our findings thus contribute to explaining the puzzle of low stock market participation: relatively illiterate individuals might tend to avoid financial assets not because of their aversion to risk but mostly because of their inability to digest financial concepts. Furthermore, we show that even a short training in basic financial concepts can foster financial literacy and increase the value assigned to financial assets.[3] This result is in line with the evidence provided by Fort et al. (2016), who show how banks' information policies can effectively increase financial literacy and, in turn, the amount of financial assets held by investors. Our design involves a purposely simple binary lottery, whereas financial products are much more complex in reality. Any effect detected in our setting is therefore likely to be a lower bound for the impact that financial literacy would have in a more complex and realistic environment.

The experimental findings can be rationalized by ambiguity aversion. Illiterate

[3]Brugiavini et al. (2015) and Lührmann et al. (2015) also show that short-term courses can significantly improve young adults' understanding of financial concepts.

agents face more ambiguity when making their investment decisions and therefore discount the value of financial assets if they are ambiguity averse. Thus, financial literacy can reduce the ambiguity faced at the time of the investment choice and increase investors' willingness to hold financial assets. In this sense, our paper is related to Dimmock et al. (2016), who show how ambiguity aversion can explain low stock market participation and the tendency of US investors not to own stocks. The authors also show that the negative effect of ambiguity aversion is stronger for illiterate agents, consistent with the findings from our experiment.[4] A model of ambiguity aversion and asset valuation – such as the one developed by Maccheroni et al. (2013) – in which the ambiguity faced by an investor decreases with her level of financial literacy, explains our main results as well as the positive link between individual comprehension of the lottery and its certainty equivalent.

This paper mainly contributes to the empirical literature focusing on the relationship between financial literacy – or the lack thereof – and economic behavior. Lusardi and Mitchell (2014) provide an exhaustive summary of this vast body of literature. Financial literacy is found to be closely linked to several economic outcomes. It tends to be highly correlated with the activity in financial markets (Christelis et al., 2010; van Rooij et al., 2011; Christelis et al., 2011)[5] as well as with the likelihood of undertaking retirement planning (Lusardi and Mitchell, 2007, 2011). Additionally, financially literate households are more capable to face negative macro-shocks (Klapper et al., 2013) and benefit from significantly higher returns on their savings accounts (Deuflhard et al., 2018). A lack of financial knowledge is systematically correlated with many inefficient financial habits: illiterate agents underdiversify their portfolios and rebalance them less frequently;[6] they pay higher mortgage costs (Moore, 2003), fees and transaction costs and are more likely to use high-cost methods of borrowing (Lusardi and de Bassa Scheresberg, 2013; Lusardi and Tufano, 2015).[7]

Despite the abundance of empirical studies, few of them provide causal evidence of the role of financial literacy, mostly because of the possible endogeneity

[4]Other related works on ambiguity aversion and stock market participation include those by Easley and O'Hara (2010) and Izhakian and Yermack (2017).

[5]Similar evidence is also provided by Kimball and Shumway (2006). Additionally, Almenberg and Dreber (2015) show that women's lack of financial literacy might explain the gender gap in stock market participation. Relatedly, Cole et al. (2014) finds a causal effect of education, in general, on stock market participation.

[6]See Gaudecker (2015), Guiso and Jappelli (2009), Calvet et al. (2007) and Calvet et al. (2009).

[7]Campbell (2006) also documents that less-educated households tend to make financial mistakes more often than their more educated counterparts; Brown et al. (2016) provide evidence that financial education improves the debt behavior of young Americans. Relatedly, Agarwal et al. (2015) show that making card-holders aware of the savings on interests that they could achieve by paying of their balance faster improves their repayment behavior, even though the effect is relatively small.

of financial literacy to financial behavior. Both financial literacy and financial behavior are likely to be jointly affected by other variables that are not always observable – for instance, cognitive abilities or social class status – and whose omission can bias the estimated coefficients.[8] Additionally, financial literacy might be determined itself by financial behavior since frequent activity in financial markets might lead to higher levels of financial sophistication. Starting from Christiansen et al. (2008) – who use new university openings as an instrument for financial education – several authors address the endogeneity issue by applying instrumental variable techniques.[9] However, this stream of literature is likely to suffer from the scarcity of non-weak, truly exogenous instruments for financial literacy.[10] Alternative approaches to IV estimations – mostly randomized field experiments – deliver mixed evidence. Some authors document that enhancing financial literacy can have an impact on saving and borrowing decisions (Sayinzoga et al., 2016; Haliassos et al., 2017), boost retirement planning (Song, 2015; Duflo and Saez, 2003), improve financial behavior (Drexler et al., 2014) and lead to higher accumulation of wealth (Bernheim et al., 2001). Conversely, other studies find no significant changes in the financial behavior of subjects exposed to financial literacy programs.[11] Carpena et al. (2018), for instance, show that financial education programs fail to improve the individual financial behaviors, unless individuals are either incentivized to set concrete financial goals after being exposed to the education treatment, or the latter is supplemented with personalized financial counseling.

The mixed evidence resulting from previous studies and the policy importance of the topic plea for further research on the linkages between financial literacy and financial behavior. Additionally, digging deeper into the channels behind this relation is crucial to understand the means through which financial literacy can affect households' financial decisions and to design truly effective financial education programs. This papers follows this direction by investigating how financial literacy influences a so-far neglected dimension of financial decision-making: the way agents value financial assets and, consequently, their propensity to acquire them. The experimental results show that improving the understanding of financial products' fundamentals and characteristics *via* a short financial literacy treatment can

[8]Theoretical foundations for the endogeneity of financial literacy come from Jappelli and Padula (2013) and Lusardi et al. (2017).

[9]See, for instance, Lusardi and Mitchell (2010), Sekita (2011), and Klapper et al. (2013) who instrument financial literacy with mandatory financial education in high schools, language ability and newspapers diffusion, respectively.

[10]In their extensive meta-analysis of the effects of financial literacy on different financial behaviors, Fernandes et al. (2014) show that some of the instruments used in previous studies either do not pass the tests for non-weak, exogenous instruments or fail to deliver significant estimates.

[11]See, for instance, Choi et al. (2011) and Collins (2013). See also Hastings et al. (2013a) for a more detailed review of the related studies.

indeed reduce households' aversion to financial products whose returns, riskiness and costs are often hard to comprehend. Endowing small investors with adequate financial knowledge can therefore significantly increase their willingness to undertake risky investments and to participate in financial markets.

The rest of the paper is organized as follows. Section 2.2 details the experimental design, and Section 2.3 describes the sample of participants. The main results are then presented in Section 2.4. Section 2.5 discusses the role of ambiguity and ambiguity aversion in explaining our findings. Finally, Section 2.6 concludes.

2.2 Experimental Design and Empirical Specification

Measuring the effect of financial literacy on individual behavior in financial markets is empirically challenging. For instance, agents who participate more in financial markets – because of their individual preferences – might end up being more financially literate than those who are less prone to do so. At the same time, individuals endowed with a greater stock of financial knowledge might be more inclined to purchase financial products than relatively illiterate agents.

We tackle this endogeneity concern by designing a randomized laboratory experiment with two randomized treatments in a two-by-two setting. We ask participants to evaluate a risky lottery, varying both the framing of the lottery and the respondents' level of financial literacy. Regarding the framing, the lottery is randomly presented either as a *simple* lottery (a coin toss) or as a *financial* lottery (a risky financial asset). In both cases the structure of the lottery – that is, the payoffs and the associated probabilities – is exactly the same. In particular, the gamble yields either 14 euros or nothing with equal probabilities. While the framing of the *simple* lottery does not involve any financial concept, a full comprehension of the *financial* one requires a few financial rudiments. Participants evaluating the financial lottery are offered a financial asset issued by a hypothetical company (AeroFlights SA). The financial asset has a current value of ten euros and yields a net return of 40% by the end of the experiment when participants get paid the final value of the asset unless the issuing company defaults. In this case, which occurs with a 50% probability, the final value is zero. The default probability is 50%. We elicit each participant's certainty equivalent of the risky lottery – presented either as a coin toss or as a financial asset – following Holt and Laury (2002). Participants make 20 sequential choices between the lottery and different safe amounts of money, ranging from 50 cents to 10 euros. Within this framework we measure the certainty equivalent for each participant as the safe amount at which she stops

accepting the lottery and switches to the safe alternative.[12]

The second treatment we provide in our two-by-two design is a teaching of basic financial notions. The teaching treatment consists of a page explaining – in a simplified and stylized way – what a financial asset is, how to calculate returns and what occurs in the case of default of the issuer.[13] It therefore induces an increase in participants' level of financial literacy. Participants are randomly assigned to the teaching treatment regardless of the way the lottery is presented and receive it immediately before making their choices. The two treatment dimensions – *i.e.*, the financial framing and the *teaching* – define four different groups that are reported in Figure 2.1. Participants in group S evaluate the *simple* lottery and do not receive *teaching*; those in group F evaluate the *financial* lottery and do not receive *teaching* either; finally, participants in groups ST and FT evaluate the *simple* and *financial* lottery, respectively, after receiving the *teaching*. The text of the lottery, as presented in each of the four groups, is presented in Appendix 2.8.

Teaching treatment \ Financial framing	0	1
0	**S**	**F**
1	**ST**	**FT**

Figure 2.1: Experimental design

Within this two-by-two setting, we can test whether the financial framing affects the propensity to undertake risk by comparing the average certainty equivalent of the lottery for participants in groups S and F. Furthermore, we can assess the role of financial literacy when evaluating financial assets by comparing the behavior of respondents in groups F and FT. This design allows us to test in two ways whether any effect of the teaching treatment is due to an actual increase in the level of financial literacy of treated participants, rather than due to any other behavioral change induced by the provision of the teaching. First, we compare the average certainty equivalent of the lottery in groups S and ST. As long as the *teaching* is increasing financial literacy without altering the behavior of participants – for instance, by providing incentives for participants to concentrate more on their choices or making them more confident – no differences should emerge in the subsample of participants evaluating the coin-toss lottery. Additionally, we

[12]More precisely, we do not identify the precise individual certainty equivalent but rather a 50-cent interval in which it lies.

[13]See Appendix **??** for details about the information provided with the *teaching*.

test whether the *teaching* actually increases the respondents' understanding of the risky lottery when financially framed. To measure both their perceived and objective understanding, we ask participants to report how much they think they understood about the structure of the lottery (on a scale from 0 to 10) and to indicate both the maximum and expected gain from the lottery. We then ask participants how useful they found the *teaching* when making their choices – on a scale from 0 to 10 – and which of the information provided with the *teaching* was the most useful.

Before the lottery choice, we present a set of ten financial literacy questions to assess the *ex ante* degree of financial sophistication of participants.[14] The questionnaire resembles the ones widely used to measure financial knowledge. It includes the standard questions about inflation, interest compounding and diversification, some more advanced questions about bonds, stocks, and options and two others requiring some numerical computations.[15] Our individual index of financial literacy equals the number of correct answers to the survey, thus ranging from 0 to 10. We use this index to measure the correlation between *ex ante* financial literacy and individuals' certainty equivalent and to check whether our sample reproduces the financial literacy patterns found in the existing literature.

Finally, at the end of the experiment, we ask participants to evaluate a risky and ambiguous lottery. Participants have to make twenty sequential choices between drawing a ball from a box containing green and blue balls in unknown proportion and a safe amount of money. When choosing the ball, participants can win either five euros if the ball is green or nothing otherwise. The safe amount offered ranges from 0.25 to 5 euros, and the certainty equivalent of this lottery is defined as the safe amount at which an agent switches from choosing the box to the safe alternative.[16] Since the lottery is the same in all groups and is shown at the end of the experiment, we compare the average certainty equivalent in the four groups to ensure that the treatments do not systematically alter respondents' attitude towards risk and ambiguity.

Given the experimental design, we infer the causal impact of financial literacy on the subjective value assigned to the financial asset by estimating the following equation:

$$CE_i = \alpha + \gamma TEACH_i + \delta FINLOT_i + \beta TEACH_i \times FINLOT_i + \phi X_i + \varepsilon_i, \tag{2.1}$$

where CE_i is the certainty equivalent of the risky lottery for individual i. $TEACH_i$

[14]Prior to the financial literacy test, we ask a few additional questions through which we collect personal information about age, income, education, etc. See Appendix 2.8 for a complete list of such questions.

[15]See Appendix 2.8 for the detailed list of the financial literacy questions used in the survey.

[16]The risky and ambiguous lottery offered to respondents is detailed at the end of Appendix 2.8.

is a dummy variable that equals 1 when respondent i receives the *teaching* – that is, the exogenous increase in her level of financial literacy. $FINLOT_i$ is an indicator that takes a value of one if individual i is given the financial lottery and zero otherwise. Finally, X_i is a vector of individual controls including, among others, gender, age, education, and income. In this specification, δ measures the effect of the financial framing on the certainty equivalent of the lottery. γ and β capture the impact of an increase in financial literacy on the value assigned to the coin-toss gamble and to the financial asset, respectively.[17]

The effect of financial literacy on the propensity of respondents to undertake the risky option – that is, the signs of β and γ – is *a priori* ambiguous. On the one hand, agents lacking financial literacy might be more averse to undertaking the risky lottery when it is presented as a financial asset (negative δ). On the other hand, less financially sophisticated agents might overestimate the value of the financial asset (positive δ) if they perceive it as less risky than the coin toss. In both cases, financial literacy can mitigate (amplify) the effect of the framing, in which case β would have the opposite (same) sign of δ. We expect γ to be statistically no different from zero: as long as the *teaching* affects participants' behavior only through an increase in financial literacy, the *teaching* should only be effective in the subsample of agents evaluating the *financial* lottery.

We then test whether the *teaching* effectively improves participants' understanding of the lottery by estimating the following equation:

$$UND_i = \alpha + \gamma TEACH_i + \delta FINLOT_i + \beta TEACH_i \times FINLOT_i + \phi X_i + \varepsilon_i, \tag{2.2}$$

where UND_i is either the self-assessed understanding of the lottery – how much participant i believes she understood about the structure of the lottery – or the objective understanding. In the latter case, UND_i is a dummy variable that equals 1 when the participant is able to compute the maximum and average gain achievable when choosing the risky option.[18] Thus, the δ coefficient identifies the difference in the understanding of the lottery due to the financial framing, and β measures whether – and how much – the *teaching* is effective in mitigating this distortion. Once again, we expect the γ coefficient to be statistically no different from zero.

[17] We treat the outcome variable as continuous – being grouped in 20 small-sized intervals – and estimate Equation (2.1) using OLS, thus facilitating the interpretation of the estimated coefficients.

[18] More specifically, we use three different measures of objective understanding: i) an indicator variable that equals one if the participant can correctly compute the maximum win from the risky lottery; ii) an indicator variable that equals one if the participant can correctly calculate the average win; and iii) a dummy variable that equals one only if the participant is able to correctly report both.

2.3 Sample Description

The experiment was run at the Behavioral Sciences Laboratory (*BESLab*) of University Pompeu Fabra in Barcelona in December 2016. All participants were recruited via an E-mail invitation sent to all the subjects in the database of the *BESLab*. In total, eleven sessions with approximately 24 participants each occurred in a computerized classroom over two days. Each session lasted approximately 50 minutes including payment. Subjects' earnings ranged between 5 (the show-up fee) and 24 euros, with an average of 14.62 euros.

Our sample is composed of 260 participants randomly divided into four groups. Table 1.1 presents descriptive statistics of the respondents. Approximately 65% of participants are female, and their age ranges from 18 to 41, the average being 21. Only 24% of participants took a finance class before participating in the experiment, and approximately one-third of them studied either economics and finance or political sciences. When asked to assess their own level of financial knowledge on a scale from 0 to 10, participants reported an average level of 4. The average score in the financial literacy test equals 5.5. Table 2.1 also reports the differences in means of all of these variables and other individual characteristics among the different groups. No statistically significant differences emerge, confirming that groups are balanced across all the observed dimensions.

Figure 2.2 provides a graphical representation of the distribution of the financial literacy measure in the sample. The median score in the financial literacy test is 5. Approximately 20% of respondents correctly answer fewer than 4 questions, and only 5% score 10 out of 10. The share of correct answers is widely heterogeneous across questions. More than 70% of participants correctly answer to the questions about inflation and diversification and are able to define what a stock is. Conversely, less than 30% of them know about the relationship between bond prices and interest rates. Additionally, no more than 40% of respondents are able to compute the expected value of a simple scratch card and the value of a property in two years, given the yearly percentage increase in its price. Between 50% and 60% of correct answers were collected on the remaining questions about interest rates, riskiness of stocks *vs.* bonds, call options and bond definition. These results are largely comparable with evidence from similar surveys in other countries. For instance, van Rooij et al. (2011) also find that only 24.6% of respondents in the 2005-2006 DNB Household Survey correctly answer the question about interest rate and bond prices (29% in our sample), while 63.3% know about diversification (70% in our case); 60.2% recognize that stocks are normally riskier than bonds (56.92% in our survey), and 55.5% know what a bond is (in our case, 58.08%).

By regressing our measure of *ex ante* financial literacy on the respondents' characteristics, we document that our sample replicates all of the usual patterns

of financial literacy found by most of the previous studies in the literature.[19] The results of these regressions are reported in Table 2.2. The financial literacy score is significantly lower for females and tends to increase with family income (even though this effect vanishes when controlling for education). As one would expect, participants with a degree in economics and finance, as well as participants who took a finance course during their careers, score higher than the rest of the respondents. Similar results emerge when using self-assessed financial literacy as the dependent variable.

2.4 Results

The randomized nature of the experiment allows us to estimate the main effects of our treatments by simply comparing the average certainty equivalent in the four groups. Figure 2.3 provides a graphical representation of the two main findings. First, the financial framing reduces the average value assigned to the risky lottery by 20%. Second, participants exposed to the coin-toss lottery do not change their behavior when receiving the *teaching*, while the latter effectively enhances the subjective value of the lottery when it is framed as a financial asset.

Table 2.3 presents the main results from a systematic analysis of these effects. In particular, the table contains the estimated coefficients of Equation (2.1) under different specifications. In all of the regressions, the dependent variable is the certainty equivalent of the lottery, evaluated as the safe amount at which an individual starts preferring the safe amount to the risky alternative.[20] Column (1) presents the OLS estimate with neither controls nor fixed effects. In Columns (2) to (4), we include additional individual controls and session fixed effects. Finally, Column (5) reports the estimates from a Tobit model that accounts for the upper limit on the safe amount offered to participants, which is equal to ten euros.

The estimated coefficients presented in Table 2.3 confirm that both the financial framing and the teaching treatment have a sizable impact on the choice made by the individuals in our sample. In particular, receiving the *teaching* increases the value assigned to the financial asset by around 30% (+1.3 to +1.5 euros, depending on the specification). This increase completely offsets the negative effect due to the financial framing (between -0.93 and -1.09 euros). The sum of the two

[19] See Lusardi and Mitchell (2014) for an exhaustive review of these works and a summary of their results.

[20] We exclude from the analysis those individuals switching more than once in our setup *à la* Holt and Laury (2002), for whom we cannot observe a unique certainty equivalent. In our sample, these are approximately 25% of the observations. Importantly, the number of multiple switchers in the four groups is not systematically different, as shown in Table 2.1.

coefficients is indeed no statistically different from zero.[21] We do not find any evidence of an effect of the *teaching* on the behavior of agents facing the coin toss, as the estimated coefficient of $TEACH$ is zero. Consistent with the hypothesis that the value assigned to the financial lottery depends on financial literacy, we also find that both measures of pre-treatment financial literacy – the score in the test and the self-assessed one – are associated with higher certainty equivalents (even though the coefficient of the self-assessed measure is not significant). Lastly, it is worth noting that the Tobit coefficients do not differ from the OLS ones since most of the participants choose a switching point that is strictly included in the interval made available to them.

In Table 2.4, we also estimate the effect of the *teaching* on the probability of choosing the lottery for each of the twenty safe amounts offered. We run twenty different regressions – one for each amount proposed – where the dependent variable is a binary indicator that equals one if an individual accepts the lottery. We present the estimates obtained when including multiple switchers in the sample (Panel A) and when excluding them (Panel B). The results from this additional analysis show that the *teaching* is particularly effective in increasing the probability of choosing the risky option when the safe amount ranges between 5.5 and 6.5 euros. This evidence is consistent with the fact that increasing financial literacy impacts the behavior of the marginal respondent. Neither the very risk-averse nor the very risk-loving investor is significantly affected by the provision of financial sophistication. Figure 2.4 illustrates this finding by plotting the share of respondents that opt for the risky alternative for each of the twenty safe amounts offered.

To test whether the increase in financial literacy translates into a better understanding of the lottery's structure, we estimate how the random assignment to the groups affects the individual comprehension of the gamble. We consider both a subjective – on a 0 to 10 scale – and three objective measures for the level of understanding: i) $Correct_1$ is a dummy that equals 1 when the participant correctly identifies the maximum win achievable when choosing the lottery (14 euros); ii) $Correct_2$ is a dummy that equals 1 when the participant correctly infers the average win from the lottery (7 euros); and finally, $Correct$ is a dummy that equals 1 when both questions are correctly answered.

Table 2.5 presents the results from estimating Equation (3.2). Consistent with the hypothesis that agents have difficulties in comprehending the characteristics of the lottery when it is presented as a financial asset, the coefficient of $FINLOT$ is negative and largely significant. The financial framing indeed reduces the subjec-

[21] Testing the joint significance of the estimated coefficients $\hat{\beta}$ and $\hat{\delta}$ from Equation 2.1 returns an F–statistic of .63 (P–value= .43). The F–statistic from the test $\hat{\beta} + \hat{\delta} = \hat{\gamma}$ is also statistically no different from 0 (F–statistic= .04 and P–Value= .84).

tive understanding of the lottery by more than 2 points (on a scale from 0 to 10) and the probability of correctly recognizing the expected (maximum) win from the risky lottery by 54% (62%). The *teaching* significantly improves both the self-assessed and the objective understanding of the framed lottery. The coefficient of $FINLOT \times TEACH$ is positive and significant in all columns. Receiving the *teaching* when the lottery is presented as a financial asset increases the reported understanding of the lottery by 1.7 to 1.8 points, thus nearly offsetting the negative effect of the financial framing. Additionally, it increases the share of participants who are able to identify both the expected and maximum win from the lottery by 48%. Finally, the *teaching* given to agents facing the *simple* lottery does not significantly affect their understanding (none of the coefficients of $TEACH$ is significantly different from zero).[22] Hence, the *teaching* enhances financial literacy and increases individuals' ability to comprehend the payoffs and the riskiness associated with the *financial* lottery.

Additional evidence on the importance of the *teaching* comes from Figure 2.5, where we plot the distribution of the "usefulness" of the *teaching* for agents facing the *simple* and the *financial* lottery, respectively. As expected, the majority of respondents evaluating the *financial* lottery find the *teaching* particularly useful when making their choice and assign to it an average value of 6.8 out of 10, more than 3 points higher than that observed in the group of respondents evaluating the coin toss. Furthermore, when asked about which of the information provided with the *teaching* respondents find most useful, the majority of participants in the *simple* group find none of the notions of some use. In the *financial* group, the answer selected by most of the participants is "how to compute returns". Learning about returns was indeed useful when making the choice about accepting the lottery or not.[23]

Lastly, we analyze the participants' behavior when they are offered the risky and ambiguous lottery at the end of the experiment. We exploit these responses in a placebo test to check whether the provision of either of the two treatments alters agents' attitude towards risk and ambiguity. The results from this test are shown in Table 2.6. None of the coefficients is statistically different from 0, thus showing that receiving or not receiving the teaching treatment only affects individual behavior when making choices involving financial concepts and does not change individuals' underlying aversion towards risk and ambiguity.

[22]The F–statistics from testing the hypothesis $\hat{\beta} + \hat{\delta} = \hat{\gamma}$ on the estimated coefficients of Equation 3.2 when the measure of understanding considered is either the subjective understanding or $Correct$ are both statistically no different from zero. The F–statistics are 0 (P–value=.98) and .34 (P–value=.56), respectively.

[23]Figure 2.6 provides a graphical representation of this result by showing the difference in the share of participants choosing each of the four possible answers in the *simple* and *financial* groups.

2.5 Financial Literacy and Ambiguity

The experimental evidence described so far shows that i) individuals lacking financial literacy tend to discount the value of financial assets and ii) increasing financial literacy reduces this distortion, thus promoting the willingness to invest in financial products. We also show that enhancing financial literacy – through the provision of the teaching treatment – increases participants' understanding of the payoffs and risks associated with financial assets. A possible explanation for these results relies on the concept of ambiguity aversion. Indeed, financial illiteracy is likely to increase the ambiguity faced by an investor when making her investment choice, thus lowering her willingness to undertake risk and purchase financial products. Reasonably, the probability of committing a mistake when evaluating a financial asset depends on the level of financial literacy of the investor, and illiterate agents have a higher probability of either overestimating or underestimating the value of the asset. When agents are not naive, they are perfectly aware of this higher probability of committing mistakes, and if ambiguity averse, the fear of overestimating a *bad* asset outweights the chance of underestimating a *good* one. As a result, illiterate agents have lower incentives to invest in financial markets when ambiguity averse.

The important role of ambiguity and ambiguity aversion as determinants of financial behavior has been highlighted by several recent studies. For instance, Dimmock et al. (2016) document a negative association between ambiguity aversion and stock market participation in a sample of US households. Similarly, Izhakian and Yermack (2017) provide evidence that ambiguity tends to increase early exercising of executives' stock options. Additionally, theoretical models of portfolio allocation under ambiguity, such as the ones developed by Easley and O'Hara (2010) and Maccheroni et al. (2013), show that investors discount the value of ambiguous assets.

Hence, our experimental results can be interpreted in this light. Consider, for instance, a decision maker who evaluates an ambiguous prospect according to the *smooth* model of decision under ambiguity proposed by Klibanoff et al. (2005). Maccheroni et al. (2013) show that the analogous of the Arrow-Pratt approximation for the certainty equivalent in the presence of ambiguity would be:

$$C_i(h) = E_P(h) - \frac{\lambda_i}{2}\sigma_P^2(h) - \frac{\theta_i}{2}\sigma_\mu^2(E(h)), \tag{2.3}$$

where λ_i measures the decision maker's absolute risk aversion; θ_i is a parameter capturing her degree of ambiguity aversion; $\sigma_P^2(h)$ is a measure of the risk implied by the stochastic nature of the prospect h; and finally, $\sigma_\mu^2(E(h))$ is a measure of the ambiguity faced by the decision maker, that is, the uncertainty regarding the

true model according to which h is distributed.[24]

This model delivers a simple and tractable approximation for the value that an agent assigns to a risky and ambiguous payoff, which can be easily brought into our experimental setting and allows us to link financial literacy to the value that an ambiguity averse agent assigns to a financial asset. Indeed, one can think that an individual evaluating a risky asset faces two sources of uncertainty: the first one is the physical uncertainty $\sigma_P^2(h)$, that is, pure *risk*, driven by the variability in the possible realizations of the payoff; the second one is model uncertainty $\sigma_\mu^2(E(h))$, that is, the *ambiguity* in the true probabilistic model according to which the final value of the asset is distributed. In this framework, financial literacy can impact the subjective value of financial assets by reducing the ambiguity the agent faces – that is, $\sigma_\mu^2(E(h))$ – and therefore increase the certainty equivalent of a risky financial asset for an ambiguity averse investor. In a nutshell, financial literacy increases the value that an ambiguity averse decision maker assigns to a risky asset since it enhances her ability to correctly identify the fundamentals of the asset, thus reducing the ambiguity she faces during the valuation process.

Consistent with this mechanism, we show that enhancing financial literacy through the provision of the teaching treatment increases the participants' understanding of the payoffs and risks associated with the *financial* lottery and makes it more valuable to participants. A positive association between the individual comprehension of the gamble and its certainty equivalent also emerges when focusing on the group of participants who evaluate the *financial* lottery and do not receive the *teaching*. Exploiting the across-subject heterogeneity in the comprehension of the *financial* lottery's structure, we find that respondents who report a higher understanding – or correctly answer the questions on the maximum and average win achievable with the lottery – indeed assign a higher value to the financial asset.

Table 2.7 presents this additional evidence. Estimates in Columns (1) and (2) show that a one-unit increase in the self-assessed understanding scale is associated with an increase of approximately 30 cents in the certainty equivalent of the financial lottery. Furthermore, the estimates in Columns (3) and (4) document that agents able to compute the maximum (average) win from the lottery value it 2.29 (2.37) euros more than those who are not able to do so. Finally, correctly inferring both the maximum and average gain from the financial lottery increases the average certainty equivalent by 3.41 euros. Hence, understanding the structure of the lottery makes it significantly more valuable to participants.

[24]See Maccheroni et al. (2013) and Klibanoff et al. (2005) for details on the model of smooth ambiguity aversion and the derivation of robust mean-variance preferences.

2.6 Conclusions

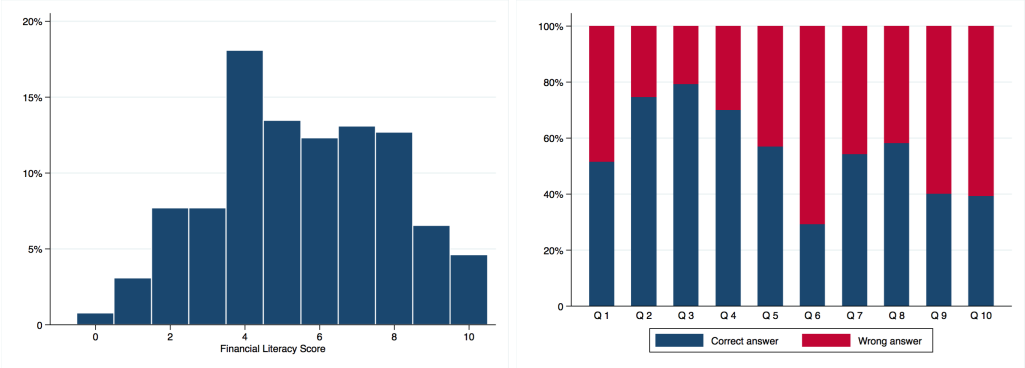
Financial literacy potentially affects financial behavior in many ways. Providing causal evidence of these effects is challenging since financial literacy is likely to be endogenous to several individual characteristics that determine financial behavior. Moreover, little is known about the potential channels through which financial literacy may affect households' financial choices, including the decision to participate in financial markets. In this paper, we contribute to fill this gap by presenting results from a laboratory experiment designed to test whether an exogenous increase in financial literacy – induced by explaining basic financial notions – affects the subjective valuation of risky financial assets.

Our results show that financial literacy plays an important role at the time of the investment decision. Framing the lottery with financial concepts, rather than presenting it as a simple coin toss, makes it more obscure to participants and reduces their willingness to take it. However, agents experiencing an exogenous increase in their level of financial literacy during the experiment value the financial lottery more than their untreated counterparts. Additionally, they report a higher understanding of the structure of the financially framed lottery. These results can be rationalized by ambiguity aversion: enhancing financial literacy can indeed reduce the ambiguity faced by retail investors when evaluating financial assets, thus increasing the subjective value of the assets.

These results have important policy implications. Our findings show that financial illiteracy may lead to a systematic undervaluation of financial products because of the lack of an adequate comprehension of the payoffs and risks associated with the products. Therefore, endowing households with the tools to cut through the fog of financial markets – even through a concise training of relevant financial concepts – can have a sizable impact on their willingness to invest in financial products and participate in financial markets.

2.7 Figures and Tables

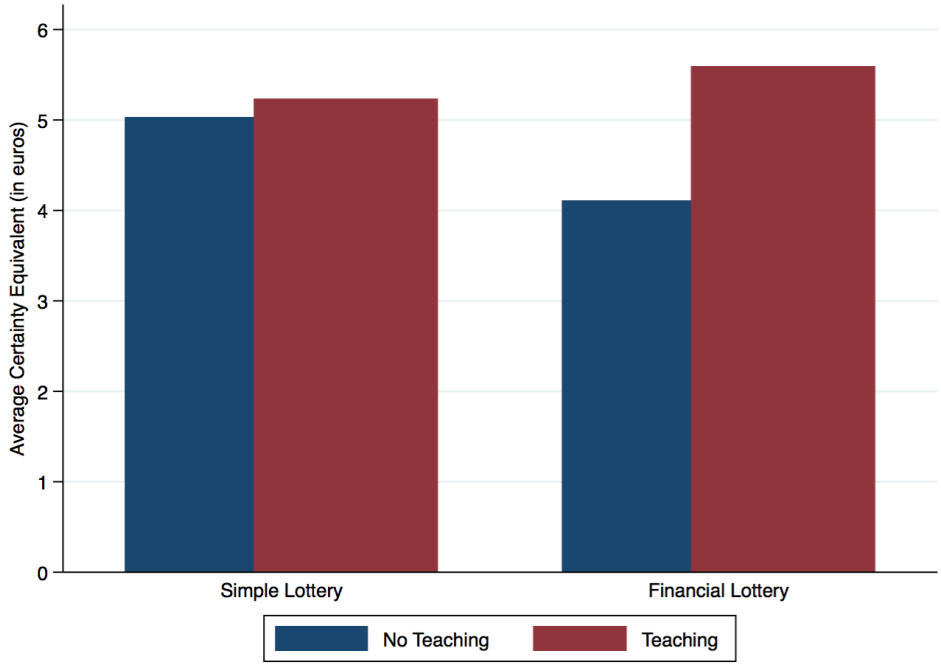
Figure 2.2: Distribution of pre-treatment financial literacy



Aggregate By question

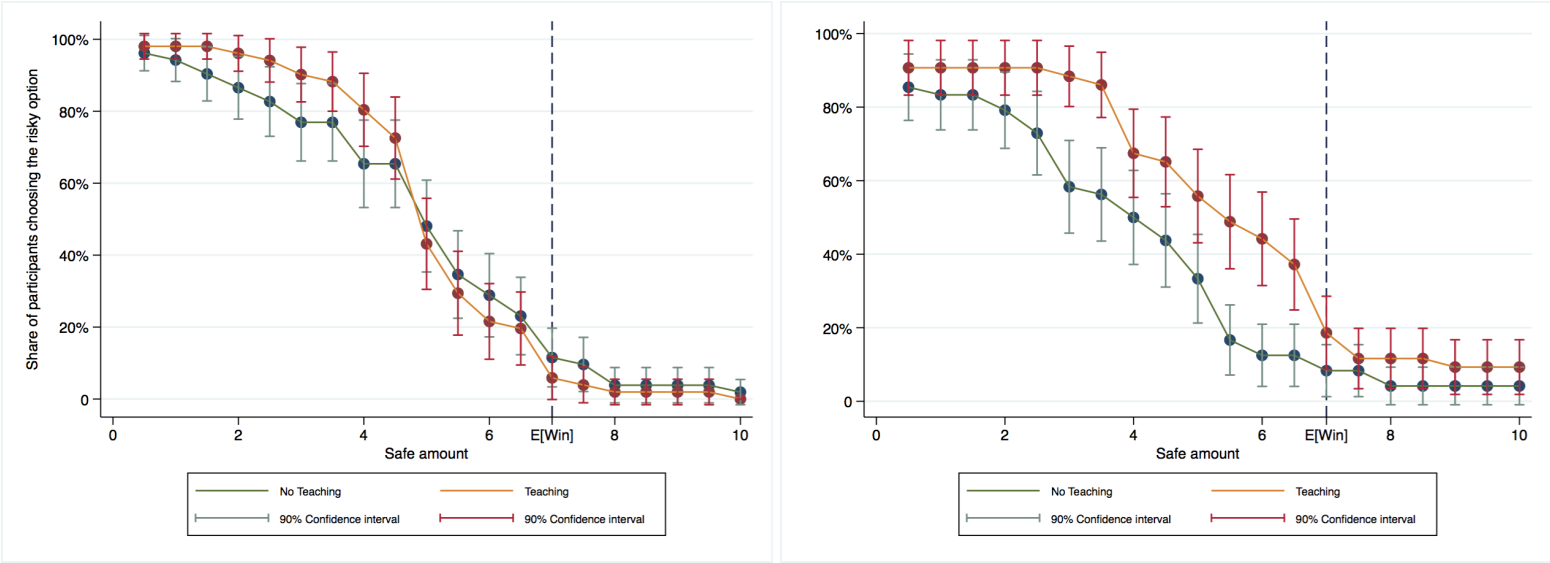
Note: The left panel of this figure plots the distribution of the score obtained in the financial literacy test. The score equals the number of correct answers and goes from 0 to 10. The right panel details the share of correct and wrong answers to each of the ten questions in the test.

Figure 2.3: Average certainty equivalent



Note: This figure plots the average certainty equivalent (in euros) of the risky lottery in each of the four groups. The certainty equivalent corresponds to the safe amount at which respondents stop choosing the risky option and switch to the safe alternative.

Figure 2.4: Heterogeneity in the probability of taking the risky lottery

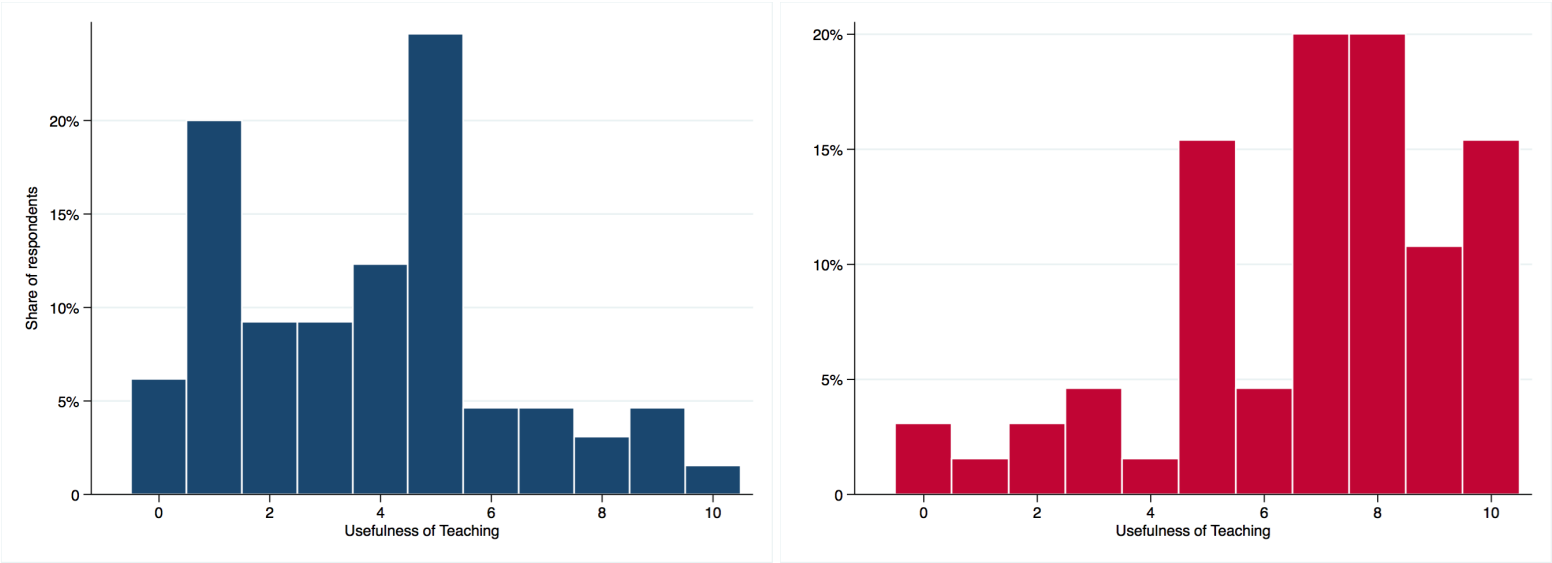


Simple lottery

Financial lottery

Note: This figure plots the probability of taking the risky lottery for each of the twenty safe amounts offered. The left (right) panel plots the share of respondents – receiving and not receiving the teaching treatment in orange e green, respectively – who take lottery when this is presented as a coin toss (financial asset).

Figure 2.5: Usefulness of the teaching treatment (1)

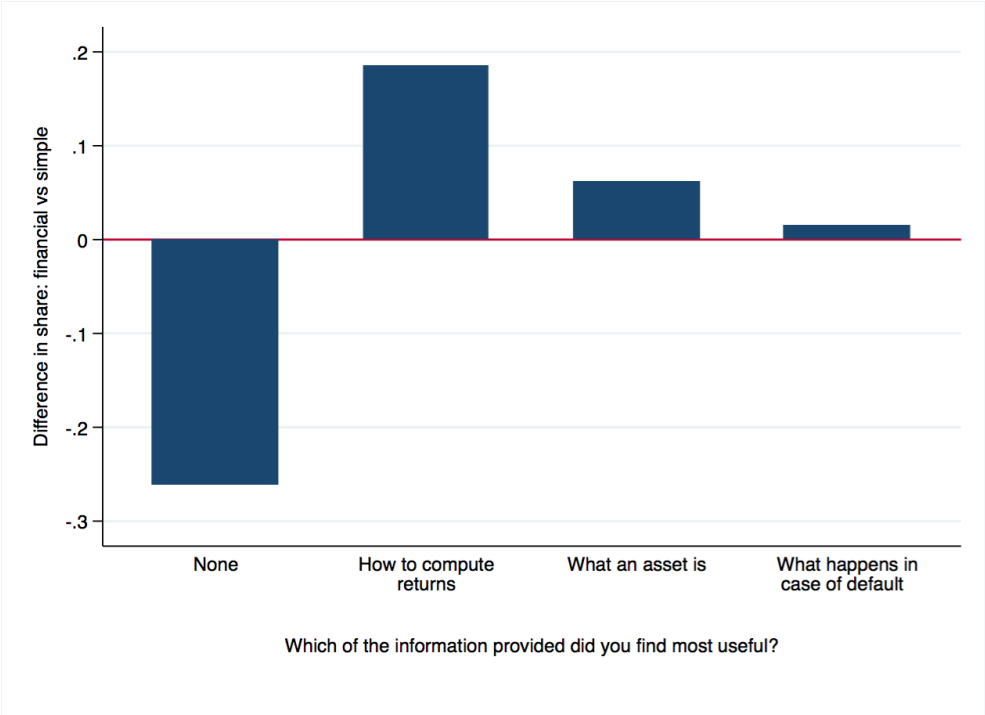


Simple lottery

Financial lottery

Note: This figure plots the distribution of the value – on a scale from 0 to 10 – assigned to the usefulness of the *teaching* by participants in the *simple*-lottery (left panel) and in the *financial*-lottery group (right panel).

Figure 2.6: Usefulness of the teaching treatment (2)



Note: This figure plots the difference in the share of respondents between the *financial*-lottery and the *simple*-lottery groups who indicated each of the four possible answers to the question *"Which of the information provided did you find most useful when making your choices?"*. Negative values correspond the a larger share of respondents picking that option in the *simple*-lottery group, whilst positive values result from a larger share of participants choosing that answer in the *financial*-lottery group.

Table 2.1: Summary statistics

Participants characteristics and mean differences between groups	Mean	St. Dev.	$\mu_F - \mu_S$	$\mu_{ST} - \mu_S$	$\mu_{FT} - \mu_S$
Female	0.65	0.48	-0.06	-0.03	0.05
Age	21.14	3.29	0.35	-0.77	-0.02
Work	0.32	0.47	-0.05	-0.06	-0.02
Working years	1.48	2.47	-0.02	-0.32	-0.29
Family Income $>$ 80K euros	0.05	0.23	0.00	0.02	-0.02
Family Income 40K-80K euros	0.27	0.45	-0.03	0.00	0.14
Family Income $<$ 40K euros	0.54	0.50	0.05	0.00	-0.11
Education level: High School Diploma	0.12	0.32	-0.02	0.09	0.03
Education level: Bachelor's Degree	0.74	0.44	0.02	-0.06	-0.06
Education level: Master	0.10	0.31	0.00	-0.03	0.02
Education level: PhD	0.02	0.15	0.02	-0.02	0.00
Field of studies: Economics/Finance/Pol. Sciences	0.35	0.48	0.00	-0.03	0.06
Field of studies: Humanities/Law	0.29	0.45	0.02	0.02	-0.02
Field of studies: Medicine/Biology/Psichology	0.12	0.33	0.00	-0.05	0.02
Field of studies: Other	0.24	0.43	-0.02	0.06	-0.06
Took a finance course	0.24	0.43	0.03	-0.03	0.05
Self-assessed Financial Literacy (0-10)	4.24	1.81	0.05	-0.54	-0.25
Financial Literacy score (0-10)	5.53	2.37	0.03	-0.08	-0.14
Multiple Switchers	0.25	0.44	0.06	-0.02	0.14
C.E. Ambiguous Lottery	2.68	1.06	0.12	0.17	0.10

NOTE: This table reports the summary statistics of the whole sample of participants, as well as the difference between the mean of each variable in group *S* (simple lottery with no *teaching*) and in the three other groups. *, ** and *** indicate that the mean difference is statistically different from 0 at the 99, 95 and 90% confidence level, respectively. *Financial Literacy score* is the score obtained in the financial literacy test, while *Self-assessed Financial Literacy* is the individual self-assessed level of financial knowledge. Both measures are on a 0 to 10 scale. Except for *Age* and *Working years* all of the other variables are binary indicators. *Multiple Switcher* is a dummy variable that equals 1 if we cannot uniquely identify, in the Holt and Laury (2002) setup, the certainty equivalent of the risky lottery for a participant.

Table 2.2: Financial literacy patterns

Pre-treatment Financial Literacy	(1) Financial Literacy	(2) Financial Literacy	(3) Financial Literacy	(4) Self-assessed Fin. Lit.
Female	-1.189***	-1.150***	-0.899***	-0.578***
	(0.300)	(0.302)	(0.241)	(0.218)
Age	0.027	0.057	0.087	-0.074
	(0.043)	(0.068)	(0.062)	(0.056)
Work		-0.312	-0.244	-0.228
		(0.337)	(0.273)	(0.248)
Working years		-0.017	0.026	0.142**
		(0.092)	(0.077)	(0.069)
Family Income > 80K euros		0.615	0.497	1.082**
		(0.643)	(0.512)	(0.461)
Education level: Bachelor's Degree			0.492	0.453
			(0.351)	(0.317)
Education level: Master			0.645	0.856*
			(0.526)	(0.474)
Education level: PhD			0.274	0.454
			(0.914)	(0.823)
Field of studies: Economics/Finance/Pol. Sciences			2.735***	0.604**
			(0.321)	(0.289)
Field of studies: Humanities/Law			0.539*	-0.047
			(0.318)	(0.289)
Field of studies: Medicine/Biology/Psichology			0.613	-0.286
			(0.412)	(0.371)
Took a finance course			1.211***	1.062***
			(0.294)	(0.265)
Constant	5.733***	5.159***	2.385*	5.126***
	(0.957)	(1.366)	(1.240)	(1.117)
Mean Dep. Var.	5.531	5.531	5.531	4.236
Standard dev.	2.365	2.365	2.365	1.806
Observations	260	260	260	258
R^2	0.060	0.067	0.436	0.217

NOTE: This table reports the estimates of a set of regressions of the measures of pre-treatment financial literacy on participants' characteristics. The dependent variable in Columns (1) to (3) is the score obtained in the financial literacy test, while the dependent variable in Column (4) is the the individual self-assessed level of financial knowledge. Both measures are on a scale from 0 to 10. Except for *Age* and *Working years* all of the other variables are binary indicators. Standard errors in parenthesis. *** $p< 0.01$, ** $p< 0.05$, *$p< 0.10$.

Table 2.3: Main results

Dependent Variable: Certainty equivalent of the risky lottery					
	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	OLS	OLS	Tobit
FINLOT	-0.925**	-1.096**	-1.062**	-1.088**	-1.091**
	(0.435)	(0.449)	(0.440)	(0.446)	(0.423)
TEACH	0.206	0.156	0.161	-0.021	0.143
	(0.429)	(0.440)	(0.431)	(0.446)	(0.415)
FINLOT$\times TEACH$	1.282**	1.405**	1.398**	1.566**	1.458**
	(0.626)	(0.652)	(0.639)	(0.651)	(0.615)
Financial Literacy score (0-10)			0.259***		
			(0.095)		
Self-assessed Financial Literacy (0-10)				0.117	
				(0.102)	
Constant	5.029***	7.659***	6.786***	6.891***	8.215***
	(0.302)	(2.291)	(2.269)	(2.323)	(2.215)
Session FE	No	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	Yes
Observations	194	194	194	192	194
R-squared	0.059	0.266	0.299	0.272	

NOTE: This table reports the estimates of Equation (2.1). The dependent variable in all columns is the certainty equivalent of the risky lottery, defined as the safe amount at which an individual starts preferring the safe alternative to the lottery. *TEACH* and *FINLOT* are dummy variables that equal 1 if an individual receives the teaching treatment and the financial framing, respectively. *Financial Literacy score* is the score obtained in the financial literacy test, while *Self-assessed Financial Literacy* is the individual self-assessed level of financial knowledge. Both measures are on a 0 to 10 scale. Controls in Columns (2) to (5) include *Female*, *Age*, *Work*, and a set of binary indicators for a participant's level of education, field of study and family-income class. In Columns (2) to (5), we also include session fixed effects: a set of dummy variables identifying the different experimental rounds. In all regressions, we exclude those individuals for which we cannot uniquely identify, in the Holt and Laury (2002) setup, the certainty equivalent of the risky lottery (*Multiple Switchers*). Estimates in Columns (1) to (4) are from OLS regressions, while a Tobit model – which accounts for the upper limit on the safe amount offered to participants (10 euros) – is assumed in Column (5). Standard errors in parenthesis. *** p< 0.01, ** p< 0.05, *p< 0.10.

Table 2.4: Effect heterogeneity, by safe amount offered

Dependent Variable: Taking the risky lottery																				
Panel A: including *Multiple switchers*																				
	0.50	1.00	1.50	2.00	2.50	3.00	3.50	4.00	4.50	5.00	5.50	6.00	6.50	7.00	7.50	8.00	8.50	9.00	9.50	10.00
FINLOT	-0.138***	-0.077	-0.185***	-0.138**	-0.062	-0.169**	-0.123	-0.169**	-0.200**	-0.154*	-0.138	-0.138*	-0.062	-0.046	0.000	0.108*	0.046	0.046	0.062	0.062
	(0.051)	(0.058)	(0.057)	(0.067)	(0.065)	(0.074)	(0.075)	(0.083)	(0.084)	(0.086)	(0.085)	(0.078)	(0.080)	(0.064)	(0.064)	(0.059)	(0.056)	(0.055)	(0.056)	(0.051)
TEACH	0.015	0.046	0.062	0.092	0.108*	0.138*	0.169**	0.123	0.092	-0.123	-0.046	-0.062	0.000	-0.046	-0.031	0.062	-0.031	-0.000	-0.000	-0.015
	(0.051)	(0.058)	(0.057)	(0.067)	(0.065)	(0.074)	(0.075)	(0.083)	(0.084)	(0.086)	(0.085)	(0.078)	(0.080)	(0.064)	(0.064)	(0.059)	(0.056)	(0.055)	(0.056)	(0.051)
$FIN \times TEACH$	0.046	-0.015	0.092	0.077	-0.015	0.154	0.092	0.092	0.108	0.308**	0.262**	0.354***	0.154	0.215**	0.138	-0.000	0.062	0.108	0.031	0.200***
	(0.072)	(0.082)	(0.081)	(0.095)	(0.092)	(0.104)	(0.106)	(0.118)	(0.118)	(0.122)	(0.120)	(0.111)	(0.114)	(0.090)	(0.090)	(0.083)	(0.079)	(0.077)	(0.079)	(0.072)
Constant	0.954***	0.892***	0.908***	0.815***	0.815***	0.723***	0.692***	0.631***	0.646***	0.492***	0.415***	0.308***	0.292***	0.154***	0.138***	0.046	0.092**	0.062	0.077*	0.031
	(0.036)	(0.041)	(0.041)	(0.048)	(0.046)	(0.052)	(0.053)	(0.059)	(0.059)	(0.061)	(0.060)	(0.055)	(0.057)	(0.045)	(0.045)	(0.041)	(0.040)	(0.039)	(0.040)	(0.036)
Observations	260	260	260	260	260	260	260	260	260	260	260	260	260	260	260	260	260	260	260	260
R-squared	0.044	0.020	0.072	0.047	0.027	0.080	0.070	0.049	0.048	0.025	0.026	0.055	0.014	0.035	0.021	0.034	0.017	0.040	0.016	0.116
Panel B: excluding *Multiple switchers*																				
	0.50	1.00	1.50	2.00	2.50	3.00	3.50	4.00	4.50	5.00	5.50	6.00	6.50	7.00	7.50	8.00	8.50	9.00	9.50	10.00
FINLOT	-0.107**	-0.109**	-0.071	-0.074	-0.098	-0.186**	-0.207**	-0.154	-0.216**	-0.147	-0.179*	-0.163*	-0.106	-0.032	-0.013	0.003	0.003	0.003	0.003	0.022
	(0.051)	(0.055)	(0.058)	(0.064)	(0.070)	(0.079)	(0.081)	(0.093)	(0.096)	(0.099)	(0.092)	(0.086)	(0.083)	(0.062)	(0.055)	(0.044)	(0.044)	(0.042)	(0.042)	(0.037)
TEACH	0.019	0.038	0.077	0.095	0.114*	0.133*	0.113	0.150	0.072	-0.049	-0.052	-0.073	-0.035	-0.057	-0.057	-0.019	-0.019	-0.019	-0.019	-0.019
	(0.051)	(0.054)	(0.057)	(0.063)	(0.069)	(0.078)	(0.080)	(0.092)	(0.094)	(0.098)	(0.090)	(0.085)	(0.082)	(0.061)	(0.054)	(0.043)	(0.043)	(0.042)	(0.042)	(0.037)
$FIN \times TEACH$	0.034	0.036	-0.003	0.020	0.064	0.168	0.185	0.024	0.142	0.274*	0.374***	0.390***	0.282**	0.159*	0.090	0.093	0.093	0.070	0.070	0.071
	(0.074)	(0.078)	(0.083)	(0.092)	(0.101)	(0.114)	(0.117)	(0.134)	(0.138)	(0.143)	(0.132)	(0.124)	(0.119)	(0.089)	(0.080)	(0.064)	(0.064)	(0.061)	(0.061)	(0.053)
Constant	0.962***	0.942***	0.904***	0.865***	0.827***	0.769***	0.769***	0.654***	0.654***	0.481***	0.346***	0.288***	0.231***	0.115***	0.096**	0.038	0.038	0.038	0.038	0.019
	(0.036)	(0.038)	(0.040)	(0.044)	(0.049)	(0.055)	(0.056)	(0.065)	(0.066)	(0.069)	(0.063)	(0.060)	(0.057)	(0.043)	(0.038)	(0.031)	(0.031)	(0.029)	(0.029)	(0.026)
Observations	194	194	194	194	194	194	194	194	194	194	194	194	194	194	194	194	194	194	194	194
R-squared	0.037	0.039	0.033	0.037	0.052	0.094	0.089	0.053	0.050	0.025	0.057	0.065	0.043	0.022	0.011	0.026	0.026	0.016	0.016	0.033

NOTE: This table reports the estimates from a set of regressions of the probability of taking the lottery, for each of the twenty safe amounts offered, on receiving the teaching treatment and the financial framing. In each column, the dependent variable is a binary indicator that equals 1 if an individual takes the lottery when offered the safe amount reported in the column header (ranging from 0.50 to 10 euros). *TEACH* and *FINLOT* are dummy variables that equal 1 if an individual receives the teaching treatment and the financial framing, respectively. In Panel A, we include those individuals for which we cannot uniquely identify, in the Holt and Laury (2002) setup, the certainty equivalent of the risky lottery (*Multiple Switchers*). They are excluded in Panel B. Estimates in all columns are from OLS regressions. Standard errors in parenthesis. *** $p < 0.01$, ** $p < 0.05$, *$p < 0.10$.

Table 2.5: Effects on the understanding of the lottery

Dependent Variable: Individual Comprehension of the Lottery					
	(1)	(2)	(3)	(4)	(5)
	Underst.	Underst.	Correct_1	Correct_2	Correct
FINLOT	-2.115***	-2.310***	-0.535***	-0.616***	-0.610***
	(0.416)	(0.401)	(0.090)	(0.086)	(0.087)
TEACH	-0.454	-0.644	-0.001	0.026	-0.105
	(0.410)	(0.392)	(0.088)	(0.085)	(0.085)
FINLOT$\times TEACH$	1.680***	1.826***	0.247*	0.563***	0.477***
	(0.599)	(0.582)	(0.131)	(0.125)	(0.126)
Financial Literacy score (0-10)		0.070	0.040**	0.038**	0.048**
		(0.087)	(0.019)	(0.019)	(0.019)
Constant	8.865***	7.859***	0.977**	0.764*	0.757*
	(0.289)	(2.064)	(0.464)	(0.445)	(0.447)
Session FE	No	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	Yes
Observations	194	194	194	194	194
R-squared	0.130	0.413	0.418	0.428	0.440

NOTE: This table reports the estimates of Equation (3.2). The dependent variable is the individual comprehension of the structure of the risky lottery, either self-assessed on a 0 to 10 scale – in Columns (1) and (2) – or objectively measured – in Columns (3) to (5). $Correct_1$ is a dummy that equals 1 when the participant correctly identifies the maximum win achievable when choosing the lottery (14 euros), $Correct_2$ is a dummy that equals 1 when the participant correctly infers the average win from the lottery (7 euros), and $Correct$ is a dummy that equals 1 when both questions are correctly answered. *TEACH* and *FINLOT* are dummy variables that equal 1 if an individual receives the teaching treatment and the financial framing, respectively. *Financial Literacy score* is the score obtained in the financial literacy test, on a 0 to 10 scale. Controls in Columns (2) to (5), include *Female*, *Age*, *Work*, and a set of binary indicators for a participant's level of education, field of study and family-income class. In Columns (2) to (5) we also include a set of binary indicators for the different experimental rounds. In all regressions, we exclude those individuals for which we cannot uniquely identify, in the Holt and Laury (2002) setup, the certainty equivalent of the risky lottery (*Multiple Switchers*). Estimates in all columns are from OLS regressions. Standard errors in parenthesis. *** p< 0.01, ** p< 0.05, *p< 0.10.

Table 2.6: Placebo test

Dependent Variable: Certainty equivalent of the risky and ambiguous lottery					
	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	OLS	OLS	Tobit
FINLOT	0.043	-0.029	-0.029	-0.022	-0.003
	(0.215)	(0.236)	(0.236)	(0.233)	(0.220)
TEACH	-0.174	-0.121	-0.122	-0.236	-0.118
	(0.213)	(0.231)	(0.232)	(0.233)	(0.216)
FINLOT$\times TEACH$	0.175	0.153	0.153	0.256	0.125
	(0.310)	(0.342)	(0.343)	(0.339)	(0.319)
Constant	2.654***	2.635**	2.647**	2.144*	2.882**
	(0.148)	(1.205)	(1.223)	(1.217)	(1.142)
Session FE	No	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	Yes
Observations	191	191	191	189	191
R-squared	0.007	0.132	0.132	0.148	

NOTE: This table reports the estimates from a set of regressions of the certainty equivalent of a risky and ambiguous lottery – presented at the end of the experiment, after the main lottery choice – on receiving the teaching treatment and the financial framing. The dependent variable in all columns is the safe amount at which an individual starts preferring the safe alternative to the risky and ambiguous lottery. *TEACH* and *FINLOT* are dummy variables that equal 1 if an individual received – during the main lottery choice – the teaching treatment and the financial framing, respectively. Controls in Columns (2) to (5) include *Female*, *Age*, *Work*, and a set of binary indicators for a participant's level of education, field of study and family-income class. In Columns (2) to (5), we also include a set of binary indicators for the different experimental rounds. In all regressions, we exclude those individuals for which we cannot uniquely identify, in the Holt and Laury (2002) setup, the certainty equivalent of the risky lottery or that of the risky and ambiguous lottery (*Multiple Switchers*). Estimates in Columns (1) to (4) are from OLS regressions, while a Tobit model – which accounts for the upper limit on the safe amount offered to participants (5 euros) – is assumed in Column (5). Standard errors in parenthesis. *** p< 0.01, ** p< 0.05, *p< 0.10.

Table 2.7: Individual understanding and the subjective value of the financial asset

Dependent Variable: Certainty equivalent of the *financial* lottery					
	Self-assessed understanding		Objective understanding		
	(1)	(2)	(3)	(4)	(5)
Understanding	0.271**	0.272*			
	(0.118)	(0.144)			
Correct_1			2.294**		
			(0.939)		
Correct_2				2.373***	
				(0.871)	
Correct					3.406**
					(1.390)
Constant	2.273**	-0.082	1.250	1.036	2.222
	(0.864)	(1.998)	(2.379)	(2.346)	(2.369)
Controls	No	Yes	Yes	Yes	Yes
Observations	48	48	48	48	48
R-squared	0.103	0.300	0.338	0.359	0.338

NOTE: This table reports the estimates from a set of regressions of the certainty equivalent of the *financial* lottery on the comprehension of its structure by participants. The sample is restricted to individuals who receive the financial framing but not the teaching treatment. The dependent variable in all columns is defined as the safe amount at which an individual starts preferring the safe alternative to the financial asset. $Understanding$ is the individual self-assessed comprehension of the structure of the risky lottery, on a 0 to 10 scale. $Correct_1$ is a dummy that equals 1 when the participant correctly identifies the maximum win achievable when choosing the lottery (14 euros), $Correct_2$ is a dummy that equals 1 when the participant correctly infers the average win from the lottery (7 euros), and $Correct$ is a dummy that equals 1 when both questions are correctly answered. Controls in all columns include *Female*, *Age*, *Work*, and a set of binary indicators for a participant's family-income class. In all regressions, we exclude those individuals for which we cannot uniquely identify, in the Holt and Laury (2002) setup, the certainty equivalent of the risky lottery (*Multiple Switchers*). Estimates in all columns are from OLS regressions. Standard errors in parenthesis. *** p< 0.01, ** p< 0.05, *p< 0.10.

2.8 Appendix

PART I:
GENERAL QUESTIONS

Age:

Gender:

- Male
- Female

Highest level of education (achieved or current):

- High School diploma
- Bachelor
- Master
- PhD
- Other

Field of study (if any):

- Economics or Finance
- Accounting or Management
- Law
- Humanities
- Political Science
- Sciences or Biology
- Mathematics or Physics
- Psychology
- Medicine
- Others

Have you ever taken a Finance course during your studies?

Yes	No
○	○

Are you currently working?

Yes	No
○	○

How many years have you been working?

Approximately, what is your family net income (after taxes) per year?

- less than €20,000
- between €20,000 and €40,000
- between €40,000 and €60,000
- between €60,000 and €80,000
- between €80,000 and €100,000
- More than €100,000
- Don't know

How would you rate, on a scale from 0 to 10, your financial knowledge?

	0	1	2	3	4	5	6	7	8	9	10
Knowledge											

PART II:
FINANCIAL LITERACY TEST

Suppose you have €100 in a savings account and the interest rate is 20% per year. If you never withdraw money or interest payments, how much would you have in this account after 5 years?

More than €200	Exactly €200	Less than €200	Do not know
○	○	○	○

Imagine that the interest rate on your savings account is 1% per year and inflation is 2% per year. After 1 year, how much would you be able to buy with the money in this account?

More than today	The same as today	Less than today	Do not know
○	○	○	○

Which of the following statements is correct? If somebody buys the stock of firm B in the stock market:

- ○ He owns a part of firm B
- ○ He has lent money to firm B
- ○ He is liable for firm B's debts
- ○ None of the above;
- ○ Do not know

When an investor spreads his money among different assets, the risk of losing money:

Increases	Decreases	Stays the same	Do not know
○	○	○	○

Stocks are normally riskier than bonds. True or false?

True	False	Do not know
○	○	○

If the interest rate falls, what should happen to bond prices?

Rise	Fall	Stay the same	Do not know
○	○	○	○

When you buy a Call option on a stock, you are actually buying:

- ○ The right to sell a stock at a certain price in the future
- ○ The right to buy a stock at a certain price in the future
- ○ The obligation to sell a stock at a certain price in the future
- ○ The obligation to buy a stock at a certain price in the future
- ○ Do not know

Which of the following statements is correct? If somebody buys a bond of firm B:

- He owns a part of firm B
- He has lent money to firm B
- He is liable for firm B's debts
- None of the above;
- Do not know

Someone gives you a scratch card that allows you to win:

- €10 with probability 1/2
- €16 with probability 1/4
- €20 with probability 1/4

Compute and indicate the expected payout. If you do not know, write "Do not know".

Example: if your answer is €10, write:
"10"

If the value of an apartment increases by 5% per year and today it is worth €450,000, how much will it be worth in two years?

Indicate your answer below, in euros. If you do not know, write "Do not know".

Ex: if your answer is €20000, write:
"20000"

Group S: *Simple* lottery with no *teaching*

PART III:
DECISION-MAKING

By completing the first two parts of the experiment you will be paid 5 euros.

Now you have the chance to earn an additional amount of money by choosing among different options in the next questions.

We offer you:

- **a safe amount of money; or**
- **the possibility of tossing a coin.**

If you opt for the coin toss, you will receive €14 if you get **head**, and €0 if you get **tail**.

You must make 20 sequential choices between tossing the coin and earning a safe amount of money. We propose you 20 possible amounts, from €0.50 to €10, as shown in the table below.

At the end of the experiment a row among the 20 will be randomly selected and your earnings will depend on the option you selected in that row. If you had chosen the coin toss, at the end of the experiment the computer will simulate the coin toss and you will be paid according to the **outcome (head or tail)**.

Example:

in the first row, we offer you €0.50. Would you prefer the €0.50 (the safe amount) or the coin toss? And in the second row, we offer you €1, would you prefer tossing the coin or getting €1 for sure? And so on...

(**YOU MUST MAKE A CHOICE IN EACH ROW!**)

	The safe amount (on the left)	Tossing the coin
€0.50	○	○
€1.00	○	○
€1.50	○	○
€2.00	○	○
€2.50	○	○
€3.00	○	○
€3.50	○	○
€4.00	○	○
€4.50	○	○

	The safe amount (on the left)	Tossing the coin
€5.00	○	○
€5.50	○	○
€6.00	○	○
€6.50	○	○
€7.00	○	○
€7.50	○	○
€8.00	○	○
€8.50	○	○
€9.00	○	○
€9.50	○	○
€10.00	○	○

How confident are you, on a scale from 0 to 10 (where 0 is "I did not understand at all" and 10 is "I perfectly understood"), of having fully understood the previous question?

	0	1	2	3	4	5	6	7	8	9	10
Understanding											

How much do you think one wins, on average, when tossing the coin?

	0	2	4	6	8	10	12	14	16	18	20
€											

How much do you think one can win, at most, when tossing the coin?

	0	2	4	6	8	10	12	14	16	18	20
€											

Group F: *Financial* lottery with no *teaching*

PART III:
DECISION-MAKING

By completing the first two parts of the experiment you will be paid 5 euros.

Now you have the chance to earn an additional amount of money by choosing among different options in the next questions.

We offer you:

- **a safe amount of money; or**
- **a risky financial asset issued by the company AeroFlights SA.**

The financial asset has a **current value** of €10 and, with 50% probability, it will yield a **net return** of 40% at the end of the experiment. With the remaining 50% probability, AeroFlights SA will **default** and the value of the financial asset will be €0.

You must make 20 sequential choices between the financial asset and earning a safe amount of money.
We propose you 20 possible amounts, from €0.50 to €10, as shown in the table below.

At the end of the experiment a row among the 20 will be randomly selected and your earnings will depend on the option you selected in that row. If you had chosen the financial asset, you will get its **future value** (at the end of the experiment) that will be established by a market simulator according to the afore-stated probabilities.

Example:

in the first row, we offer you €0.50. Would you prefer the €0.50 (the safe amount) or the financial asset? And in the second row, we offer you €1, would you prefer the financial asset or getting €1 for sure? And so on...

(**YOU MUST MAKE A CHOICE IN EACH ROW!**)

	The safe amount (on the left)	The financial asset
€0.50	○	○
€1.00	○	○
€1.50	○	○
€2.00	○	○
€2.50	○	○
€3.00	○	○
€3.50	○	○

	The safe amount (on the left)	The financial asset
€4.00	○	○
€4.50	○	○
€5.00	○	○
€5.50	○	○
€6.00	○	○
€6.50	○	○
€7.00	○	○
€7.50	○	○
€8.00	○	○
€8.50	○	○
€9.00	○	○
€9.50	○	○
€10.00	○	○

How confident are you, on a scale from 0 to 10 (where 0 is "I did not understand at all" and 10 is "I perfectly understood"), of having fully understood the previous question?

	0	1	2	3	4	5	6	7	8	9	10
Understanding											

How much do you think one wins, on average, when choosing the financial asset?

	0	2	4	6	8	10	12	14	16	18	20
€											

How much do you think one can win, at most, when choosing the financial asset?

	0	2	4	6	8	10	12	14	16	18	20
€											

Group ST: *Simple* lottery with *teaching*

PART III:
DECISION-MAKING

By completing the first two parts of the experiment you will be paid 5 euros.

Now you have the chance to earn an additional amount of money by choosing among different options in the next questions.

Before making your choices, please open the file "AdditionalInformation.pdf" by clicking here. In this file, you will find information that might be relevant and that might help when taking your choices. Please read them carefully!

We offer you:

- **a safe amount of money; or**
- **the possibility of tossing a coin.**

If you opt for the coin toss, you will receive €14 if you get **head**, and €0 if you get **tail**.

You must make 20 sequential choices between tossing the coin and earning a safe amount of money. We propose you 20 possible amounts, from €0.50 to €10, as shown in the table below.

At the end of the experiment a row among the 20 will be randomly selected and your earnings will depend on the option you selected in that row. If you had chosen the coin toss, at the end of the experiment the computer will simulate the coin toss and you will be paid according to the **outcome (head or tail)**.

Example:

in the first row, we offer you €0.50. Would you prefer the €0.50 (the safe amount) or the coin toss? And in the second row, we offer you €1, would you prefer tossing the coin or getting €1 for sure? And so on...

(**YOU MUST MAKE A CHOICE IN EACH ROW!**)

PS: Did you remember to open the file "AdditionalInformation.pdf"? If you think it contains information useful for your decisions and you want to re-open it, here it is.

[Note of the authors: the table is the same as in group S. For brevity, we omit it here.]

How confident are you, on a scale from 0 to 10 (where 0 is "I did not understand at all" and 10 is "I perfectly understood"), of having fully understood the previous question?

	0	1	2	3	4	5	6	7	8	9	10
Understanding											

How much do you think one wins, on average, when tossing the coin?

	0	2	4	6	8	10	12	14	16	18	20
€											

How much do you think one can win, at most, when tossing the coin?

	0	2	4	6	8	10	12	14	16	18	20
€											

How useful did you find, from 0 to 10, where 0 is “useless” and 10 is “crucial”, the information provided (what a financial asset is, how to compute returns...) for your decisions?

	0	1	2	3	4	5	6	7	8	9	10
Usefulness											

Which of the information provided did you find most useful when making your choices?

- ◯ What a financial asset is
- ◯ How to compute returns/future value
- ◯ What happens when the company issuing the financial asset defaults
- ◯ None of them

Group FT: *Financial* lottery with *teaching*

PART III:
DECISION-MAKING

By completing the first two parts of the experiment you will be paid 5 euros.

Now you have the chance to earn an additional amount of money by choosing among different options in the next questions.

Before making your choices, please open the file "AdditionalInformation.pdf" by clicking here. In this file, you will find information that might be relevant and that might help when taking your choices. Please read them carefully!

We offer you:

- **a safe amount of money; or**
- **a risky financial asset issued by the company AeroFlights SA.**

The financial asset has a **current value** of €10 and, with 50% probability, it will yield a **net return** of 40% at the end of the experiment. With the remaining 50% probability, AeroFlights SA will **default** and the value of the financial asset will be €0.

You must make 20 sequential choices between the financial asset and earning a safe amount of money.
We propose you 20 possible amounts, from €0.50 to €10, as shown in the table below.

At the end of the experiment a row among the 20 will be randomly selected and your earnings will depend on the option you selected in that row. If you had chosen the financial asset, you will get its **future value** (at the end of the experiment) that will be established by a market simulator according to the afore-stated probabilities.

Example:

in the first row, we offer you €0.50. Would you prefer the €0.50 (the safe amount) or the financial asset? And in the second row, we offer you €1, would you prefer the financial asset or getting €1 for sure? And so on...

(**YOU MUST MAKE A CHOICE IN EACH ROW!)**

PS: Did you remember to open the file "AdditionalInformation.pdf"? If you think it contains information useful for your decisions and you want to re-open it, here it is.

[Note of the authors: the table is the same as in group F. For brevity, we omit it here.]

How confident are you, on a scale from 0 to 10 (where 0 is "I did not understand at all" and 10 is "I perfectly understood"), of having fully understood the previous question?

	0	1	2	3	4	5	6	7	8	9	10
Understanding											

How much do you think one wins, on average, when choosing the financial asset?

	0	2	4	6	8	10	12	14	16	18	20
€											

How much do you think one can win, at most, when choosing the financial asset?

	0	2	4	6	8	10	12	14	16	18	20
€											

How useful did you find, from 0 to 10, where 0 is “useless” and 10 is “crucial”, the information provided (what an asset is, how to compute returns...) for your decisions?

	0	1	2	3	4	5	6	7	8	9	10
Usefulness											

Which of the information provided did you find most useful when making your choices?

- ◯ What a financial asset is
- ◯ How to compute returns/future value
- ◯ What happens when the company issuing the financial asset defaults
- ◯ None of them

Ambiguous Lottery

Finally, we present you the last question, where you have the chance to earn some extra money.

We offer you:

- a safe amount of money; or
- the possibility of drawing a ball from a box containing 10 balls;

The box contains green and blue balls in **unknown proportions**. If you draw a **green** ball from the box, you earn €5, if you draw a **blue** ball from the box, your get €0.

You must make 20 successive choices between drawing a ball or earning a safe amount of money. We propose you 20 possible amounts, from €0.25 to €5, as shown in the table below.

At the end of the experiment a row among the 20 will be randomly selected and your earnings will depend on the option you selected in that row. If you had chosen to draw a ball, at the end of the experiment the computer will simulate the draw, and you will be paid according to the **color** of the ball you get (green or blue).

Example:

in the first row, we offer you €0.25. Would you prefer the €0.25 (the safe amount) or drawing a ball? And in the second row, we offer you €0.50, would you prefer drawing a ball or getting €0.50 for sure? And so on…

(**YOU MUST MAKE A CHOICE IN EACH ROW!**)

	The Safe Amount (on the left)	Drawing a ball
€ 0.25	○	○
€ 0.50	○	○
€ 0.75	○	○
€ 1	○	○
€ 1.25	○	○
€ 1.50	○	○
€ 1.75	○	○
€ 2	○	○
€ 2.25	○	○
€ 2.50	○	○
€ 2.75	○	○
€ 3	○	○
€ 3.25	○	○
€ 3.50	○	○
€ 3.75	○	○
€ 4	○	○
€ 4.25	○	○
€ 4.50	○	○
€ 4.75	○	○
€ 5	○	○

The Teaching

1. What is a financial asset?

A financial asset is a financial instrument, as for instance a stock or a bond, that can be traded in financial markets and whose value depends on the characteristics of the issuing company.

2. How do you compute the future value of a financial asset, given the rate of return?

The **future value** of a financial asset can be determined knowing its **rate of return**.

By multiplying the rate of return by the current value, you will know the return of the asset, that is, the increase in its value over time.

Therefore, the **future value** of a financial asset will simply be the sum of its current value plus the return.

Here you have a brief example: If an asset has a current value of 1000 euros, and its rate of return is 30%, the return will be: $1000 \times 30\% = 300$ euros. The future value of the asset will be $1000 + 300 = 1300$ euros.

However, this will happen only if the company does not fail.

3. What happens if the issuer company defaults?

When the issuer company defaults, the financial asset will loose all of its value and therefore its future value will be **zero**.

Chapter 3

STUDENTS' CHOICES AND THE EMPLOYMENT PREMIUM OF UNIVERSITY QUALITY

3.1 Introduction

Whether attending prestigious universities generates better wage and employment opportunities constitutes a key research topic which has been widely investigated in the economic literature. However, data limitation and selection issues – *ex ante* good students are often paired with top institutions – make the identification of such causal link a serious challenge. This paper seeks to provide empirical evidence on the existence of an employment premium to education quality. Moreover, it also tackles a key policy question, that is, whether policy-makers should explicitly push students towards selective or prestigious universities they would not have chosen otherwise.

I exploit an unexpected budget cut in a scholarship program implemented by an Italian region – Sardinia – that targeted postgraduate students who are willing to pursue their studies in other Italian regions or abroad. The *Master and Back* scholarships are awarded on the basis of both applicants' curriculum vitae – evaluated following a precise scoring grid – and the prestige of the chosen university – measured through the *Quacquarelli-Symonds University Rankings*. While in the early years of the program nearly all applications were successful, in 2010 an unexpected budget tightening made the university choice a key determinant of the probability of winning the scholarship. The implementation of

a more competitive selection process and the weight assigned to the prestige of the chosen university *de facto* constrain applicants' choice set by limiting their possibility of opting for low-ranked schools under the scholarship program. This constraint is more binding for students with a poorer curriculum as they need more points from the university choice to meet the (expected) threshold for winning the scholarship. The identification strategy relies on two sources of variation with a difference-in-differences design. I compare two cohorts of students who take their decision about which university to attend before and after the budget cut is revealed. Moreover, I exploit the heterogeneity in treatment intensity mechanically induced by the evaluation criteria. Applicants with better curricula are freer to choose the university and the city they like the most; conversely, those lacking points have fewer options, as low-ranked universities would not guarantee good chances of a successful application.

I find that the increased degree of competition for scholarships effectively push students towards higher-ranked universities. On average, MAB applicants who start a master in the academic year 2011/2012 choose institutions whose score is on average 8 points higher – out of 60 – than the average from the previous year. This effect is heterogeneous, as students belonging to the middle of the cv score distribution – namely, those in the second tercile – are the ones who increase the most their university score. Difference-in-differences estimates show that applicants with a middle-cv score, who receive the strongest incentives, are 30 percentage points more likely to be employed around 15 months after the end of their master's degree. Although I cannot distinguish between the different channels behind the observed large employment effect, I show that incentives make students switching from universities located in Central or Southern Italy to universities in the North, the latter being the area in Italy with the lowest unemployment rate.[1] Again, this effect is heterogeneous with respect to the cv score distribution, with students in the second tercile being the ones exhibiting the strongest shift. The observed increase in employment might therefore be at least partially driven by students having access to a more favorable labor market right after the end of the master .

There is a vast economic literature on the labour market returns of education quality. While the identification strategy of most of the early studies on this topic relies on OLS estimates, a number of authors in recent years emphasized the need of implementing more sophisticated approaches to address the potential endogenous sorting of students across colleges (Black and Smith 2006). Among others, Hoekstra (2009), Hastings et al. (2013b), and Saavedra (2009) exploit the discontinuous jump in the probability of enrolling in a selective college induced by ad-

[1] According to the National Statistic Institute (ISTAT) the youth unemployment rate in 2012 was 31.2% in Northern Italy, while 39.8% and 51.6% in Central and Southern Italy, respectively.

mission cutoffs. They all find that barely accepted students, when compared with those who just miss the admission threshold, benefit from a sizable and significant earning premium later in life. Alternative identification strategies are those proposed by Behrman et al. (1996) and Lindahl and Regner (2005), who use sample of twins, and by Dale and Krueger (2002) and Dale and Krueger (2014) who estimate selection-adjusted models.[2]

There might be different mechanisms underlying the positive effect of college prestige on labor market outcomes, going from human capital accumulation to peer effects and signaling in the job market. Lang and Siniver (2011) compare students from two university in lsrael – one more prestigious than the other – and suggest that the signaling channel is likely to play a major role. MacLeod et al. (2017) also highlight the importance of college reputation in the sorting process regulating how students choose their college, and find their first job.[3] Lastly, similar evidence has been provided by Hershbein (2013), who shows that grades are important - or relatively more important - for those students who are coming from institutions which name alone provides a less clear signal about students' quality.

Regression discontinuity techniques represent the most popular approach to provide causal evidence on the effect of attending selective or prestigious universities. However, analyzing the labor market performance of students whose test score is just below or above the admission threshold set by top institutions suffers from two main limitations. First, it is challenging to interpret the discontinuity parameter, as those students who miss the minimum score have potentially hundreds of different alternatives – including the option of not attending any college – which are often difficult to track (Rodríguez et al. 2016). This concern is particularly serious as most studies only look at a small set of institutions within a country (Canaan and Mouganie 2018), thus considering very few available alternatives to the most selective one. As a consequence, the counterfactual group – those just below the admission score of a selective university – is widely heterogeneous. Second, focusing on the set of students around the cutoff implies comparing the 'worse' students in a elite universities with those who are likely to be in the upper tail of the distribution in a less selective one. As both treated and control individuals might not constitute the representative student in each of the two universities,

[2]Other studies include Andrews et al. (2016). For the case of Italy, Brunello and Cappellari (2008) investigated the role of specific universities in affecting employment and earnings three years after graduation. Their result suggests that this seems to be the case, especially for private universities.

[3]More precisely, they exploit the introduction of a college exit exam in Colombia as a natural experiment in which employers benefit from a new information about potential employees' ability, thus becoming less dependent from the signal of college reputation. Since the correlation between college prestige and earnings lowers once employers can better observe students' skills, their results emphasize the signaling function of college identity in a context of limited information about individual characteristics.

results are difficult to generalize to a broader population. Lastly, both Zimmerman (2014) and Canaan and Mouganie (2018) highlight how most RD papers on the returns of college quality focus on students with relatively strong academic background. This latter study is the only one addressing – at least partially – both limitations since the authors look at the enrollment decisions of low-skilled students while observing the set of all available options within the country (France). More precisely, they analyze a mandatory examination for students in the last year of high-school, finding that marginally successful students are more likely to attend higher-quality universities – although not the most selective ones – and have higher post-degree earnings.[4]

This paper is also related to the literature on students' choices over college education and on the role of information and policy interventions. Pallais (2014) shows that students are extremely responsive to small changes in the college application procedure. A modest decrease in the application costs translates into a huge jump in the probability of applying to more institutions. Also, Cohodes and Goodman (2014) find that a scholarship program funding high-school students who opt for a public Massachusetts college has a positive effect on enrollment in such colleges. The observed change in students' enrollment decisions is surprising as the amount of the scholarship is relatively small and public in-state colleges in Massachusetts are on average of lower quality when compared to the set of available alternatives.

The contribution of the paper is two-folded. On the one hand, the findings show that there is a substantial employment premium associated with the investment in education quality. On the other hand, the framework considered – besides allowing to estimate the link between education quality and employment opportunities – is interesting *per se*. Differently from all the papers mentioned above, I study an exogenous variation in students' choice set which is induced through the provision of financial incentives. Treated students are pushed towards more prestigious universities which they would have not chosen in a fully-unconstrained scenario with more scholarships available and low pressure to maximize the application score. The results from such unique policy experiment thus suggest that incentives could make students better of in terms of employment opportunities, at least in the short run, even when they do not face particular constraints when making their choices. This evidence is consistent, among others, with the findings of Goodman (2016), who shows that US students might tend to underestimate their chance of being accepted in more selective universities, as enrollment in such colleges rises under compulsory testing.

[4]Also Zimmerman (2014) studies the effect of the reduction in admission standards, looking at students in the lower tail of the ability distribution. Crossing the minimum threshold however induces a discontinuous jump in the probability of attending *any* 4-year college, rather than on the probability of attending a more – *versus* a less – selective one.

The rest of the paper is articulated as follows: Section 3.2 discusses the characteristics of the scholarship program considered, while Section 3.4 the identification strategy. After describing the different data sources in Section 3.3, I present the first stage results – the effect of the reduction in the number of scholarships on students' choices – in Section 3.5. Section 3.6 reports the results from the analysis of the labor market premium of university quality. Finally, Section 3.7 is devoted to the discussion of the key findings and the conclusion.

3.2 *Master and Back*

The program *Master and Back* (henceforth MAB) was introduced in 2005 by the Regional Government of Sardinia – an Italian island – to encourage students to pursue postgraduate education in the rest of Italy or abroad, with the ultimate goal of strengthening the human capital of the region. The policy is composed of two parts: *i*) a scholarship program for attending a master's degree or a PhD hosted by a non-Sardinian university; *ii*) the provision of financial incentives to firms operating in the region for hiring those students who are willing to come back to Sardinia after having finished their studies. Although the two components constitute a single policy intervention, each part of MAB follows its specific rules and two separate and unrelated selection processes regulate the access to the funding. Importantly, the eligibility requirements and the selection process defined for the part *ii*) of the policy do not overlap with the ones of MAB, as they are mostly based on the characteristics of the hosting firm. I will thus focus solely on the part *i*) of the program (MAB) as detailed administrative data for the second part are not available.

The eligibility requirements for a MAB scholarship include being a Sardinian resident, holding a five-years degree with a final grade of at least 100/110, and being younger than 35 years at the time of the application. Moreover, only students with at least a conditional acceptance from the chosen university can run for the grant. Any master or PhD program – regardless of the subject – held at a university based outside Sardinia is suitable for the financial support. The Regional Administration defines five major subjects – Arts and Humanities, Engineering and Technology, Social Sciences, Natural Sciences and Life Sciences – splitting the budget unevenly among them.[5] With respect to master's degrees, two types of programs are eligible for *Master and Back*: 'standard' (first-level) masters hosted by foreign universities, and the so called 'second-level' masters offered by Italian universities. The latter are one-year programs targeted at students who already hold a 'standard' master's degree and have thus completed at

[5] Students in the majors of Engineering and Technology and Social Sciences – who also account for largest shares of applicants – have more scholarships available to them.

least five years of higher education. In spite of the fact that Italy agreed the framework implied by the Bologna Process – a 3 + 2 years university system – it is indeed still a common practice among the Italian students who hold a five-years degree to attend a further master's degree. According to the *Almalaurea* data - the most complete database on graduate students in Italy - approximately 20% of all Italian graduates pursue their studies with another master's degree, including both 'second-level' masters (7%), and standard or 'first-level' master (6%). In order to avoid confusion throughout the paper I will refer to the program chosen under MAB as "master's degree" while to the degree achieved before the application to MAB – constituting one of the eligibility requirements – as "five-years degree".

Successful applicants benefit from full or partial reimbursement of the tuition fees. Further, they receive a monthly stipend ranging between 1000 and 1500 euros depending on the family income and whether the chosen program is offered by a Italian or non-Italian university, being higher in this latter case.[6] Importantly, the rules of MAB provide for the possibility of backward funding. Since the various calls are not published regularly, students can apply for the reimbursement of the expenses for a program that they have already started or even completed. This feature of MAB implies that before choosing a postgraduate program applicants are not able to observe the rules, selection criteria and date of announcement of the specific call in which they plan to participate.

The first call of MAB was published in late 2005, later followed by six further editions, in 2007, 2008, 2009, 2010, 2011, and 2012. Figure 3.1 illustrates the timing of the different calls as well as the time window delimiting which programs are eligible for funding – depending on their start date – in each call. The comparison between the application opening period – that is the interval between the announcement of the call and its deadline – and the eligibility window of each call reveals that most of the MAB editions rely heavily on the possibility of backward funding. Students who are admitted to a university after the deadline of a MAB call must rely on the subsequent one for obtaining a scholarship.

Although the calls are not published at regular intervals, the backward-funding feature of MAB ensures that any master or PhD program started in the period 2006-2012 is suitable for funding. The different editions are however characterised by a high degree of heterogeneity in terms of selection criteria and number of scholarships available. A sharp budget tightening took place in 2010, when more than a third of the total number of submitted applications were unsuccessful. Moreover, almost half of the non-excluded applicants – that is, those fulfilling the minimum requirements – do not receive the funding. Prior to 2010, fulfilling

[6] In the 2010, 2011 and 2012 MAB calls there was a upper limit for the reimbursement tuition fees set to 12000 euros. However, this constraint is no longer binding in case the chosen university scores among the top 20 positions of the *The Times - Quacquarelli Symonds* university rankings.

the eligibility requirements often translated into obtaining the scholarship, with nearly all of the applications funded thanks to the allocation of extra resources after the publication of the call.[7] The share of successful, unsuccessful and excluded applications is reported in Figure 3.2.

MAB 2010 marked another important difference from the previous editions as a new evaluation process based on an assessment of both the applicants' cv and the prestige of the university of destination is introduced. The maximum score that an application may receive is 100 points – the maximum for the *cv* and *university* score being 40 and 60 points, respectively – with no discretion left to the officials in charge of the screening process. Applicants' curricula are evaluated according to a precise scoring grid – reported in Table 3.1 – assigning a separate score to the age of each applicant, her final grade, whether her degree was achieved within the specific time limit, and further levels of education and work experiences. While the *cv* score does not differ substantially from the ones used in the previous MAB editions, the introduction of a *university* score attached to *The Times - Quacquarelli Symonds World University Rankings* (henceforth *THE-QS* rankings) represents a major change.

The university rankings compiled by the Times Higher Education in collaboration with Quacquarelli Symonds order the higher education institutions in the world according to their academic reputation, research citations, staffing levels and internationalization. The rankings are subject-specific, and report an overall score to all of the top 300 universities in each subject.[8] The Regional Administration thus assigned the corresponding *THE-QS* score – rescaled on a range from 10 to 60 – to programs hosted by universities who are listed in the aforementioned rankings, while zero points to those who are not.[9]

3.3 Data and Sample Description

The analysis presented in the following sections is based on a unique dataset obtained by combining the administrative records relative to all MAB applicants and a quasi-census database of all Italian graduates between 2006 and 2012.

The MAB archives. The agency of the Sardinian Regional Administration in charge of MAB keeps track of all students who apply to the program. The administrative records include the baseline characteristics of each applicant (age, sex, university she graduates from, final grade, date of degree) as well as the in-

[7]This is the case, for instance, of the 2008 and 2009 editions of MAB when a later budget expansion provided enough resources for financing all of the submitted applications.

[8]The five major subjects defined by the Sardinian Regional Administration matches the ones of the *THE-QS* rankings.

[9]In each call the relevant *THE-QS* rankings are those referred to the previous year.

formation about the chosen master's degree (hosting university, name and subject of the master's degree, start and end date, tuition fee). The dataset thus reports all information which are relevant for computing both the *cv* score and the *university* score, plus the final outcome of the application.

The *Almalaurea* database. *Almalaurea* is the largest survey of Italian graduates. All students who completed a five-years degree in a university belonging to the *Almalaurea* consortium are interviewed four times: at the graduation date, and one, three, and five years after. While during the first interview students are asked about their satisfaction with their studies and their future plans, in the later interviews they are asked about their employment situation and eventual, further studies after the completion of their five-years degree. Students are surveyed via both internet-based and computer-assisted telephone interviews with an overall rate of response of approximately 80%. Moreover, *Almalaurea* collects directly from the affiliated universities each graduate's final grade, age, degree name and major subject, and demographic information.[10]

Web scraping. To complement the coverage of the *Almalaurea* database I also retrieve the publicly-available curriculum vitae of MAB applicants via web scraping. Typically, the curricula include detailed information about both previous education and professional experience, thus allowing to precisely identify the employment spells. Moreover, it is possible to determine whether the student effectively achieves the master's degree chosen under MAB or whether there is any withdrawal or termination.

I combine the unique MAB administrative records with the *Almalaurea* surveys and the publicly-available curricula of applicants. The resulting database includes a rich set of applicants' baseline characteristics, as well as their post-master employment situation. With respect to the latter, I exploit the timing of the *Almalaurea* survey. As students are interviewed several times after their five-years degree, I use the earliest interview occurred after the end of the master's degree chosen under MAB.[11] The timing of the surveys delivers at least one post-master interview for the sample of MAB participants who start a master's degree not later than three years from their five-years degree.

Table 3.2 reports the sample size and characteristics of the two cohorts of

[10]Such information are collected from the moment a university joins the *Almalaurea* consortium. The two Sardinian universities – University of Cagliari and University of Sassari – are part of *Almalaurea* since 2006. Thus, there are no data available for MAB applicants who graduate earlier than 2006.

[11]In order to be consistent with the structure of the *Almalaurea* survey I exploit the publicly-available curricula of MAB applicants to determine whether, at the time of the relevant *Almalaurea* interview, the student is employed or unemployed. Therefore, I calculate the month and the year when the *Almalaurea* interview should have happened, and use the information from the curricula to integrate and complement the *Almalaurea* surveys when a post-master interview is not available.

applicants. The number of applicants in the academic years 2010/11 and 2011/12 is 502. I exclude applicants in the field of natural sciences (35), since the number of applicants in the 2010 MAB call was lower than the number of scholarship available. Further, I drop those students who start the master later than three years from their degree (108), as the timing of the *Almalaurea* interviews does not allow to retrieve their post-master employment situation. Thus, I end up with a sample of 359 students. More than two-thirds of them are female, and they were approximately 28 years old at the master start date. The average cv score – as defined in the MAB rules – is approximately 23 points out of 40, with the strongest weight given to the final grade in the degree (on average 109/110). The share of MAB students who attend and complete the master stated in the MAB application is 75%. This number includes both those who decide to renounce before the starting date and those who drop out after enrollment. By combining the *Almalaurea* data with the publicly available professional *curricula* of MAB applicants I am able to identify the occupational status for approximately 90% of those who completed the chosen master's degree. This group of students – 242, evenly distributed among the two cohorts – constitutes the final sample for estimating the employment effects of education quality.

Table 3.2 also shows that there is substantial heterogeneity in the number of months between the master completion date and the interview. Such difference is fully driven by the date of the (pre-master) five-years degree and the master start date.[12] However, for 90% of the students the interview takes place between 5 and 24 months from the master end date. The differences between the two cohorts of applicants are discussed in Section 3.4.1.

3.4 Identification strategy

The identification strategy exploits two features of the 2010 edition of MAB: i) the number of available scholarships decreased significantly; ii) a new scoring system based also on the prestige of the chosen university is introduced. I compare the cohort of MAB applicants who enrolled in a master's degree in the academic years 2010/11 with those who did so in 2011/12, only the latter being aware of the new MAB rules and budget allocation. The 2010 MAB call was indeed published on November 21, 2010. Therefore, a large proportion of the students who started a master's degree in Fall 2010 were not able to observe the reduction in the number of scholarships available, nor the new emphasis placed on the university quality

[12]The precise month when a student is asked about her employment status by *Almalaurea* depends on the year and semester in which the she graduates. Depending on whether graduation occurs in the first or the second half of the academic year in the follow-up interviews students refer to their status as of May 1st or October 1st, respectively.

indicators. On the contrary, the following cohort of students was well informed about the new features of MAB. Furthermore, applicants in 2011/12 receive a good signal about the minimum threshold for winning the scholarship, as they are able to observe the outcome – and the rankings – of the 2010 call of MAB, which were made public in March 2011.

The total application score introduced in 2010 is a linear combination of the cv score and the university score. The choice of the university thus becomes a key determinant of the probability of receiving the scholarship. Since improving the cv score is not longer possible at the time of the application the unique option left to students who want to maximize their chances success is to opt for high-ranked universities. The new features introduced with MAB 2010 thus induce a tightening in each applicant's choice set: only universities guaranteeing enough points to overcome the (expected) thresholds can be considered as viable options for winning the scholarships. Moreover, such tightening is different depending on the endowment of cv points of each applicants. Students with a better curriculum need fewer points from the university choice and are thus less affected by the new scoring system. On the contrary, those lacking points because of a poorer curriculum face a tighter constraint and must pick a top school in order to surpass the cutoff.

I thus classify applicants belonging to the two cohorts according to their cv score, defining three groups. Students in the top tercile of the distribution (with a 'reduced' cv-score from 26 to 30) represent the 'low-treatment' group, while students in the second tercile (with a score ranging from 22 to 25) the 'high-treatment' one.[13] Whether applicants in the bottom tercile (those with a score from 0 to 21) should be classified as (more) treated or control units is *a priori* ambiguous. Although in principle these students face the largest gap to fill they are also more likely to be affected by the admission standards set by high-ranked universities. The requirements for entering top universities – based on students' grades and previous studies – indeed partially overlap the ones considered in the MAB selection process. Applicants in the bottom of the cv distribution could have low chances of being admitted to those high-ranked institutions which would have allowed them to fill the score gap with applicants with better curricula.

I thus estimate the following equation:

[13] I define the cv terciles using only the main components of the cv score, namely the final degree grade, the duration of studies, and the age at the time of the call. I do not consider points assigned to the remaining components – other experiences and further education – since these information are not available for the non-MAB students which I use as a further control group in the triple-differences strategy described below. However, as shown in Table 3.1 and Figure 3.3, the difference between the original measure and the reduced one is limited given that the excluded dimensions account at most for 10 out of 40 points.

$$y_i = \gamma_0 + \gamma_1 dT2_i + \gamma_2 d2011_i + \gamma_3 dT2_i \times d2011_i + u_i \tag{3.1}$$

with $d2011_i$ is a binary indicator taking value 1 if student i started a master between September 2011 and March 2012 and value 0 if she did so between September 2010 and March 2011, and $T2_i$ is a binary indicator taking value 1 if student i belongs to the second tercile of the cv score distribution and 0 otherwise. The dependent variable in Equation 3.1 (y_i) is either the quality of the chosen university or the post-master employment situation. For the aforementioned reasons, in the baseline specification I exclude from the regression sample students belonging to the first tercile of the cv distribution. However, I also present the estimation results obtained including this particular set of students and controlling for each applicant's cv group.

In order to address possible concerns about shocks that might have affected asymmetrically students belonging to the different cv terciles, I extend the difference-in-differences framework described above so as to include a between-regions comparison. I exploit the *Almalaurea* database to define a third control group made of those non-Sardinian students who attended a master's degree in the academic years 2010/11 and 2011/12.[14] The information included in the *Almalaurea* database allow to reconstruct the cv-score used to rank application in MAB even for non-MAB students and thus to classify them according to the above-defined cv terciles. I compare the employment outcomes of three groups of Italian students – those with a low-, middle-, and high-cv – with those of the applicants to MAB.[15]

The triple-differences equation is:

$$\begin{aligned} y = \beta_0 + \delta_1 dMAB_i + \delta_2 d2011_i + \delta_3 dT2_i + \delta_4 dMAB_i \times d2011_i \\ + \delta_5 dMAB_i \times dT2_i + \delta_6 d2011_i \times dT2_i + \delta_7 dMAB_i \times d2011_i \times dT2_i + u_i \end{aligned} \tag{3.2}$$

with $dMAB_i$ representing the variable that identifies MAB applicants in the extended sample of Italian graduated students, and $d2011_i$ is the indicator for having attended and completed a master's degree in the academic year 2011/12, rather than in 2010/11. This difference-in-difference-in-differences strategy thus

[14] Although the *Almalaurea* surveys do not include precise information about post-degree studies, such as the subject or the university, they report whether a student attended and achieved a master's degree.

[15] In the control sample I only include regions where no similar scholarship programs were in place. Moreover, I focus on regions which are well represented in the *Almalaurea* database, that is, where at least 80% of students graduate in one of the *Almalaurea* universities. Such regions are Trentino Alto-Adige, Veneto, Friuli-Venezia Giulia, Liguria, Emilia-Romagna and Umbria.

accounts for possible employment shocks which might have affected Italian students asymmetrically depending on the characteristics of their curricula. Therefore, the main identification assumption underlying Equation 3.2 is that there are no cv-group×region -specific shocks.

3.4.1 Validity of the identification strategy

The empirical strategy described in previous paragraphs relies on two main assumptions: i) the paths of the average university quality and post-master employment rate of students who attended a master before the 2010/11 academic year do not diverge between cv groups; ii) the 2010/11 and 2011/12 cohorts of MAB applicants are not different from each other.

Figure 3.9.a reports the predicted university score (Panel a) and probability of being employed after the end of the master's degree (Panel b) for each cv group over the academic years 2009/10, 2010/11 and 2011/12.[16] The groups of middle- and high-cv students seem to follow a similar trend in the pre-treatment period, confirmed by the fact that when estimating the difference-in-differences Equation 3.1 on the cohorts of applicants in 2009/10 and 2010/11 the interaction coefficient is not significant – and small in magnitude – for both the university quality (-.70, p-value 0.922) and the employment (.027, p-value 0.882) regressions. Figure 3.9.a also shows that the university choices of low-cv students evolve differently in the pre-treatment period from the ones of the other two groups. Because of this, and for the reasons mentioned in Section 3.4, I exclude this group of applicants from the main specification. As middle-cv applicants in 2011/12 are the only ones who experience an improvement in their post-master employment outcome, Figure 3.9 also suggests that university quality plays a key role in determining later labor market outcomes. Students in the middle of the cv distribution are indeed also those who increase the most their university score (see Figure 3.9.b).

I also show that the new features of MAB do not have any significant effect on the composition of pool of MAB applicants. Table 3.2 reports the comparison of the baseline characteristics of the two cohorts. For all but one of the dimensions considered there are no significant differences. Neither of the measures of students' ability – cv score, final grade in the five-years degree and completion time – differs between 2010/11 and 2011/12. Furthermore, no differences arise when looking at family income, gender and at the indicator for having earned the five-years degree from a university based outside Sardinia. The only exception – with a mean difference which is significantly different from zero at the 10% level

[16] Although there are four MAB editions prior the 2010 one data limitations do not allow to extend the analysis to earlier cohorts of applicants. The *Almalaurea* survey of graduates from Sardinian universities indeed includes only those who graduated from 2007 onwards.

– is represented by the applicants' age as of the MAB publication date. However, as the age at the master start date and the age at degree are almost identical between the two years, such difference is driven by the fact that the 2010 call was published in November 2010 while the 2011 one in June 2012. The concern that the new features of MAB might have induced – for instance – only higher-ability or richer students to apply to MAB thus finds little support in the data. Similar evidence emerges when comparing the distribution of applicants' characteristics in Figure 3.3. The distributions of applicants' cv score, age, grades, and family income are not statistically different between the two cohorts, as also shown by the Kolmogorov-Smirnov tests for the equality of distributions. Hence, taken together these results leave little space for possible threats to the identification strategy related to selection mechanisms.

Importantly, the two cohorts are equally likely to answer the post-master interview (see Table 3.2). However, applicants in 2011/12 are interviewed on average a month and a half earlier. This difference is driven, at least partially, by the fact that the master's degrees chosen under the new features of MAB have, on average, an earlier starting date. Figure 3.7 shows how there is a higher share of masters program in 2011/12 with a start date between September and December. Nevertheless, as shown in Figure 3.8, such difference is homogeneous across cv terciles. Thus, it does not represent a problem for the empirical strategy described above.[17]

As shown in Table 3.5 the new features of MAB do not affect differently master completion rate. The interaction term MAB 2011/12 $\times$ 2nd Tercile is never significant for any of the specifications reported in Columns (1) to (4). This finding thus allow to interpret the results from the employment regressions as driven by students attending better universities, rather than by differential drop-out rates.

3.5 First stage

Before discussing their labour market consequences, it is a crucial step to verify whether and how students react to the new features of MAB when choosing the master's degree. Hence, in this Section I study whether MAB applicants in 2011/12 choose higher-ranked universities and whether such effect is heterogeneous depending on the initial endowment of cv points.

Figure 3.4 shows that the share of students who apply to universities with a score of 40 points or more – that is, universities in the top 40% of the subject-specific distribution – almost doubled over the two years.[18] On average, the uni-

[17]The difference in the number of months between the master end and the interview is not significantly different across cv terciles and cohorts.

[18]This number slightly lowers when controlling for the different subjects and applicant's baseline characteristics, as shown in Columns (5) and (6) of Table 3.4.

versity score increases by more than 25% (8 point on a 0-60 scale). The estimates presented in Table 3.4 prove that this result is robust to different specifications and to the inclusion of different sets of controls. The point estimate of the 2011/2012 indicator does not vary when controlling for the different components of the cv score separately (Column 2) or when including the overall score (Column 3). Moreover, the increase in the university score is confirmed also when accounting for the left censored nature of the dependent variable with a Tobit specification.[19]

A useful insight into the interpretation of the change in students' choices comes from the area of social sciences. The most popular university among students who started a master in the academic year 2010/11 was *Sapienza University* in Rome, ranked 143th with a score of 36.17. The share of applications for a master's program in *Sapienza* fell from 21% to 5.5% in 2011/2012. Such drop is almost symmetric to the rise of the popularity of *Bocconi* – from 4.7% to 25% – which is the highest-ranked Italian university (68th) with a score of 48.67. This example is also consistent with another feature of the observed shift towards higher-ranked universities: that the reshuffle took place mostly within Italy. In Table 3.3 – presented – I report the characteristics of the master' degrees chosen by MAB applicants in 2010/11 vis-à-vis the ones chosen in 2011/12. Although incentives induce an increase in university quality we do not observe a rise in the share of master's programs in non-Italian universities. Thus, students move towards the most prestigious Italian universities for each subject. As a consequence while there is a large jump in the share of students who choose universities ranked in the top 40% of the subject-specific distribution – from 30 to 60% – still very few students go to top-10% universities, ranked 55 or above.[20]

Furthermore, Figure 3.5 – together with the estimates presented in Table 3.6, – shows that students react differently to the new features introduced with MAB 2010 depending on their initial endowment of cv points. Students in the middle of the cv score distribution are the ones implementing the strongest increase: they attend in 2011/12 programs which are ranked, on average, 12 points more than the ones chosen by their 2010/11 counterparts. Although the university scores of students belonging to the bottom and top tercile exhibit a substantial rise – 6 and 5 points, respectively – its magnitude is half of the observed increase among middle-cv students. While applicants with a good curriculum are already relatively close to the expected threshold, for those lacking cv points behaving as their colleagues

[19]Only universities above the 300th position in the *THE-QS 2010* rankings receive a positive score - from 10 to 60 points - while the others are assigned a zero score.

[20]The Italian universities with a score of 40 points or above according to the *THE - QS 2010* rankings are *Politecnico di Torino* and *Politecnico di Milano* in engineering and technology, *Sapienza University of Rome*, *University of Pisa* and *University of Bologna* in natural sciences, *Bocconi University* and *University of Bologna* in social sciences and management, *Sapienza University of Rome* and *University of Bologna* in arts and humanities.

who applied to MAB in the previous years would not have been enough to win the scholarship. Thus, we observe middle-cv students filling the gap, in terms of the total score, with their high-cv colleagues through an increase in its component linked to university quality. According to this reasoning, low-cv students are the ones who should choose the highest-ranked universities. However, they remain far from the expected cutoff score, that is, the score of the last funded application in each subject in the 2010 edition of MAB. A possible explanation involves the university admission requirements, which may act as a constraint to these students' choice set. In other words, it is likely that a poor cv – lower grades, higher completion time – harms not only students' possibilities of winning the scholarship, but also their chances of being admitted in top institutions that in turn are the only ones allowing them to fill the gap with the other applicants.

Although the incentives introduced with MAB 2010 are explicitly linked to the university rankings, they also affect the probability of moving to a certain city. In Figure 3.6 I present the predicted probability of choosing a university located in Northern Italy, Southern and Central Italy, and outside Italy separately for each of the three cv groups and for the two cohorts of applicants. The comparison of Figure 3.6.a and Figure 3.6.b confirms that the share of applicants choosing a master's degree in universities in North Italy generally increase over time, with a symmetric decrease in the preferences for universities in the rest of Italy. This effect is again heterogeneous, with middle-cv students being the ones exhibiting the largest shift towards Northern Italy universities.

A last comment is devoted to the interpretation of the magnitude of the effect across the different terciles. One might wonder why a relatively small difference in the baseline cv score – students in the second tercile have an average cv score of 24 points, while students in the third a score which is on average 4 points higher – translates into such a strong gap in the university score. However, one should keep in mind that the nature of the choice variable is, from the perspective of the student, more discrete than continuous. As quality is not the only relevant dimension in students' choice, increases in the university score might come at the expenses of a large costs in terms of distance from home. Thus, if there are strong preferences for a specific country and few universities of that country listed in the *THE* - *QS* rankings, small differences in the cv-score may require students to switch from one university to another ranked in a much higher position. Once again, this view is consistent with the evidence presented in Table 3.3, where I show that the shift happens mostly within-country rather than across-countries.

3.6 The employment premium of education quality

Postgraduate students who apply to MAB under its new features – less resources available and incentives to attend high-ranked universities – opt for master's degrees held by higher-ranked universities. In this section, I investigate whether the observed increase in university quality generates better employment opportunities.

Table 3.7 reports the estimates from the difference-in-differences Equation 3.1, where the dependent variable is the indicator for being employed in the first available post-master interview. While no significant difference between the employment rate of the two cohorts of students emerges (Column 1), Columns (2) to (4) show that students belonging to the middle of the cv distribution experience a strong and positive effect on their employment opportunities. The magnitude of the effect is large, as the estimates for the interaction term range from 0.29 to 0.32, depending on the specification. Such positive effect more than offsets the employment gap between middle- and high-cv applicants in 2010/2011, with middle-cv students being 20 percentage points less likely to be employed. Estimates are robust to extending the control group to applicants belonging to the first tercile, that is, to those who in principle have the strongest incentive to increase their university score but might fail to fulfill the admission requirements for a top university.

The triple-differences strategy described in Section 3.4 delivers similar results. Estimates are reported in Table 3.8. The coefficient of the interaction term in Column (1) suggests that the incentives introduced with MAB 2010 do not have an average positive effect on the employment of the treated cohort. However the estimates in Columns (2) to (4) – where I account for the between-terciles differences – further confirm that students belonging to the middle of the cv distribution experience a strong and positive effect on their employment outcomes. The point estimate of the triple-interaction term ranges from 0.26 to 0.27 – thus being similar to the interaction term from the double-difference model presented in Table 3.7 – and it is significant at the 10% level.

The timing of the *Almalaurea* interviews is crucial for the interpretation of the findings described above. Since students are interviewed on average approximately 15 months after the end of the master's degree, the identification strategy allows to estimate the short-run consequences of the investment in university quality, while I am not able to identify their effects in a longer horizon.

3.7 Discussion and conclusion

In this paper I study whether earning a master's degree at high-ranked universities makes students better off in terms of their post-master employment outcome.

I look at a sample of Italian postgraduate students who participate in a regional scholarship program between 2010 and 2011. I exploit both the scoring system used to rank applicants – with a component for students' curriculum vitae and one for the prestige of the chosen university – and an (unexpected) reduction in the number of funded applications. I find that applicants who are able to observe the budget cut react by choosing higher-ranked universities. Moreover, difference-in-differences estimates show that such effect is heterogeneous depending on students' cv, with students in the middle of the cv distribution – thus, far from the admission cutoff but not too much – being the ones increasing the most their university score. Importantly, I also find that the attendance rate – that is, the probability of having attended and completed the chosen master's degree – does not respond to the incentives. Students actually attend those higher-ranked universities that they have chosen under the incentive scheme. Lastly, I find that applicants who receive the strongest incentives are approximately 30 percentage points more likely to be employed around 15 months after the end of the master.

Hence, linking the applicants' score to the quality of the chosen university pushes students to apply to more prestigious institution which they would not have chosen otherwise. Besides being key for the empirical strategy for identifying the labor market returns of university quality, this result is interesting *per se*. It is indeed surprising that before the budget cut occurred students were not fully exploiting the opportunity of enrolling in a master's degree offered by a top university, as the program funded nearly all applications until November 2010. The fact itself that we observe a change in students' choices suggests that financial constraints or university admission criteria play a lesser role. The two cohorts of applicants are indeed extremely similar in terms of ability, family income and academic cv, this latter dimension being the most relevant for admission decisions. Furthermore, MAB winners benefit from a full reimbursement of tuition fees and are assigned a monthly scholarship, thus being able to afford virtually any master's program independently of their income level.[21]

Several mechanisms may underlie these findings. Students might tend to overweight the costs associated with attending a master's program offered by a more prestigious – and probably more demanding – university and to assign little weight to their future returns.[22] Similarly, students might also underestimate their chances of being accepted in selective universities (Goodman 2016), thus applying to a smaller set of institutions of lower quality. Moreover, two information mechanisms are coherent with the observed pattern of university choices. On the one hand, students may lack information about the distribution of university quality.

[21]While liquidity constraints might play a significant role, they are faced by both applicants in 2010/11 and 2011/12.

[22]For a review on the importance of information in college decisions see Bleemer and Zafar (2018).

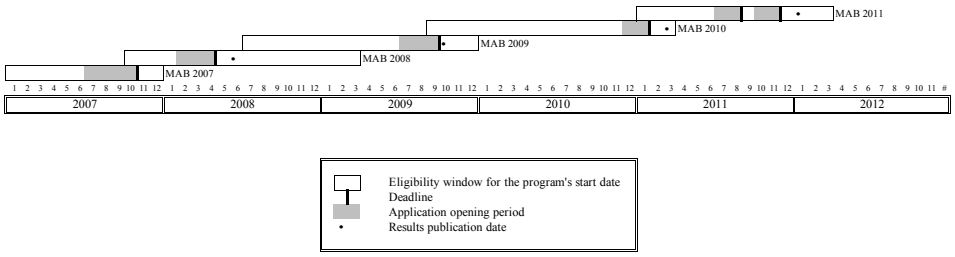
Once the competition for scholarships rises students are indeed forced to look at the *THE - QS* rankings and might learn about the reputation of each institution. On the other hand, students could also have better information – compared to the ones summarized by the *THE - QS* rankings – about the characteristics of each master's program and about its potential labor market returns. The latter being the case, linking the the probability of winning the scholarship to the rankings could harm applicants' employment prospects.

The identification strategy does not allow to fully distinguish between these alternative hypotheses. However, the results from this analysis confirm the existence of a strong link between university quality and labor market outcomes. Several authors who study the channels through which attending a prestigious university could determine better employment outcomes assign strong weight to signaling mechanisms (Lang and Siniver 2011; MacLeod et al. 2017). However, in the context under study it is worth mentioning the potential role of the specific labor market linked to each university. Among others, Oosterbeek and Webbink (2011) and Parey and Waldinger (2011) show that studying abroad is a strong determinant of living and working abroad later on in life. As the rise in the competition for the scholarships make certain universities – thus certain cities – more likely to be chosen, the observed increase in employment might be at least partially driven by students having access to a more favorable labor market. The evidence presented in Figure 3.6 suggests that incentives push students towards universities located in Northern Italy, where the unemployment rate is lower. Differences across labor markets may therefore play a key role in determining the large increase in the employment rate of middle-cv students.

Lastly, the finding of this paper have strong and novel policy implications. They suggest that policy interventions designed to reward and foster investments in education quality could be desirable even when they reduce students' choice set by eliminating low-quality options. Since students may underestimate the labor market returns of attending high-quality universities – or their chances of being admitted – they could opt for lower-quality alternatives even when not facing particular constraints. Implementing incentives to push students towards good institutions can therefore generate large labor market returns.

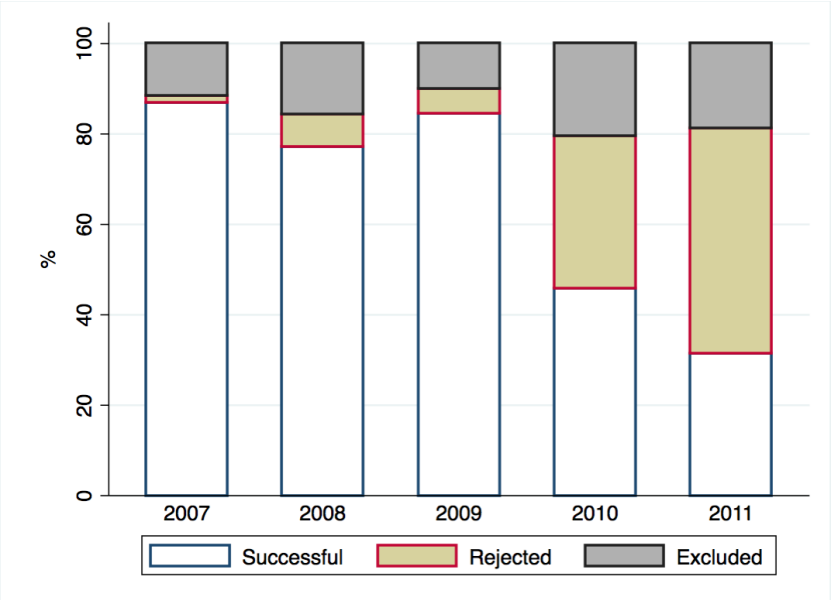
3.8 Figures and Tables

Figure 3.1: Timeline of MAB calls 2007-2011



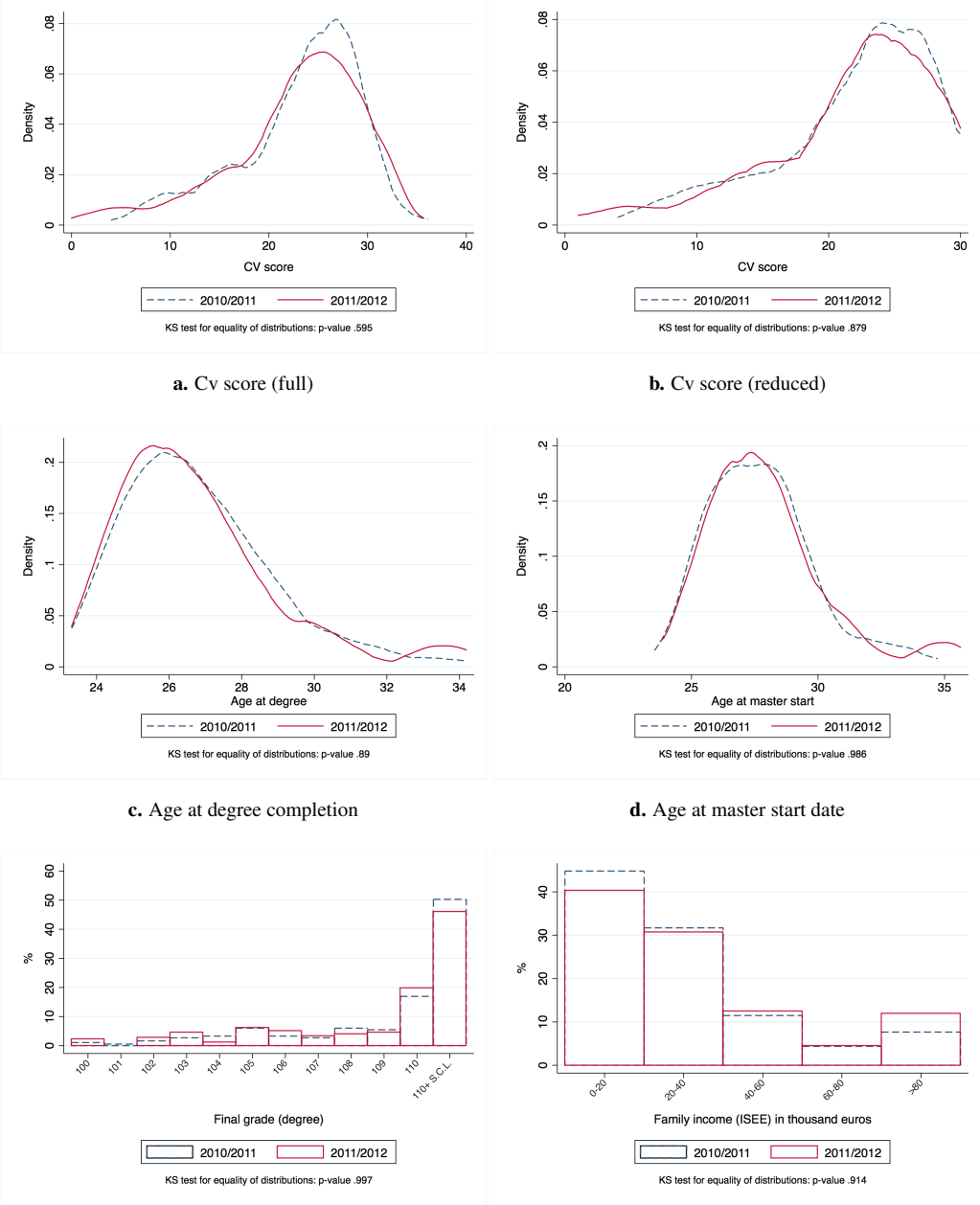
NOTES. This Figure depicts the timing of each of the MAB editions from 2007 to 2011. Each rectangle represents the time window delimiting which programs are eligible for funding – depending on their start date – in each call (*i.e. programs eligible for funding in MAB 2010 are those starting between September 2009 and March 2011.*. The thick black line in each rectangle indicates the deadline of the call, while the grey area the interval between the publication of the call and its deadline. The dot indicates the date on which results are published. From 2011 there are two selection processes within the same call, and the total number of available scholarships in a given year is split in two halves. Applicants rejected in the first half-call have the opportunity to re-apply in the second.

Figure 3.2: Share of successful applications 2007-2011



NOTES. This Figure depicts the share of successful, rejected and excluded applications in each of the MAB editions from 2007 to 2011. Only applications for a master's degree are considered. Successful applications are defined as the applications who are assigned a scholarship; rejected applications are defined as the applications which do not receive the scholarship although fulfilling the eligibility requirements; excluded applications are defined as the applications in which the student or the chosen master's degree lack one or more of the eligibility requirements.

Figure 3.3: Comparison MAB applicants 2010/2011 - 2011/2012



a. Cv score (full)

b. Cv score (reduced)

c. Age at degree completion

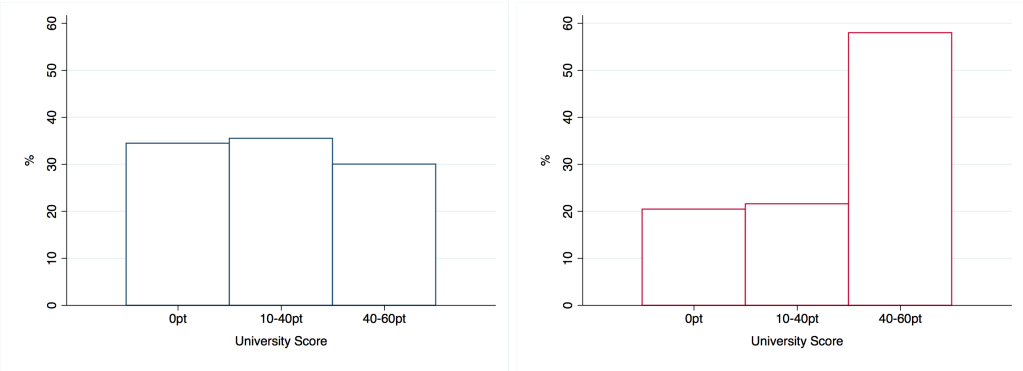
d. Age at master start date

e. Final grade (degree)

f. Family income

NOTES. This figure depicts the kernel density estimates for applicants' baseline characteristics for the two cohorts of applicants who started a master's degree in 2010/11 and 2011/12. Each Panel – from a. to f. – reports the p-value Kolmogorov-Smirnov tests for the equality of distributions. In Panel e. and f. the test is computed over the continuous distribution (not in classes) of the relevant variable. The score in panel a. and b. is the cv score computed summing all its components – reported in Table 3.1 – and excluding the points assigned to 'other experience' and 'further education', respectively. The cv terciles used throughout all the analysis are based on the latter. The sample includes applicants who have started a master's degree within two years from their five-years degree. Students in social sciences, engineering and technology, arts and humanities, life sciences and medicine are included, while students in natural sciences are excluded.

Figure 3.4: Applicants' choices: university quality



a. 2010/11 **b.** 2011/12

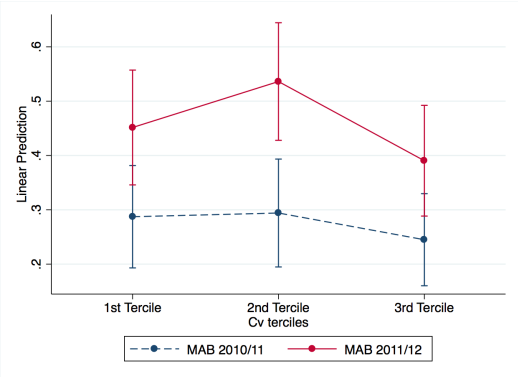
NOTES. This figure depicts the distribution of the university score for student who applied to MAB in 2010/11 (Panel a) and 2011/12 (Panel b). The sample includes applicants who have started a master's degree within two years from their five-years degree. Students in social sciences, engineering and technology, arts and humanities, life sciences and medicine are included, while students in natural sciences are excluded.

Figure 3.5: Heterogeneity of the increase in university quality

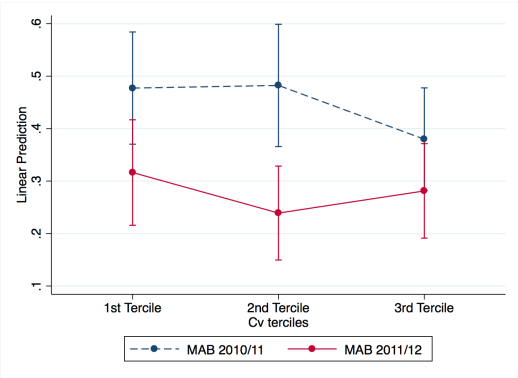
a. By cv score **b.** By cv tercile

NOTES. This figure depicts the total score of applicants in 2010/11 and 2011/12 depending on their cv score. Blue diamonds and red dots in Panel a. indicate the total score of applicants in 2010/11 and 2011/12 respectively, while the dashed blue (solid red) line is a local regression prediction of the relationship between total and cv score in 2010/11 (2011/12). The vertical distance between each value of the cv score and the dots (diamonds) thus represents each applicant's university score. Panel b. shows the relationship between cv, university and total score for the different cv terciles. Full blue dots (red diamonds) indicate the average total score in 2010/11 (2011/12), and empty blue dots (red diamonds) indicate the average cv score in 2010/11 (2011/12). Dashed blue (solid red) lines indicates the university score in 2010/11 (2011/12). In both panel the gold horizontal line indicates the average threshold score for winning the scholarship in the 2010 edition of MAB. The sample includes applicants who have started a master's degree within two years from their five-years degree. Students in social sciences, engineering and technology, arts and humanities, life sciences and medicine are included, while students in natural sciences are excluded.

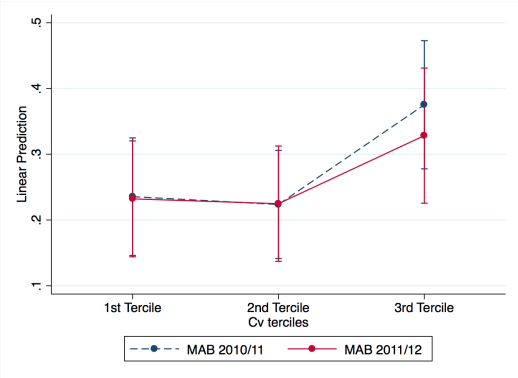
Figure 3.6: Applicants' choices: university location



a. North Italy



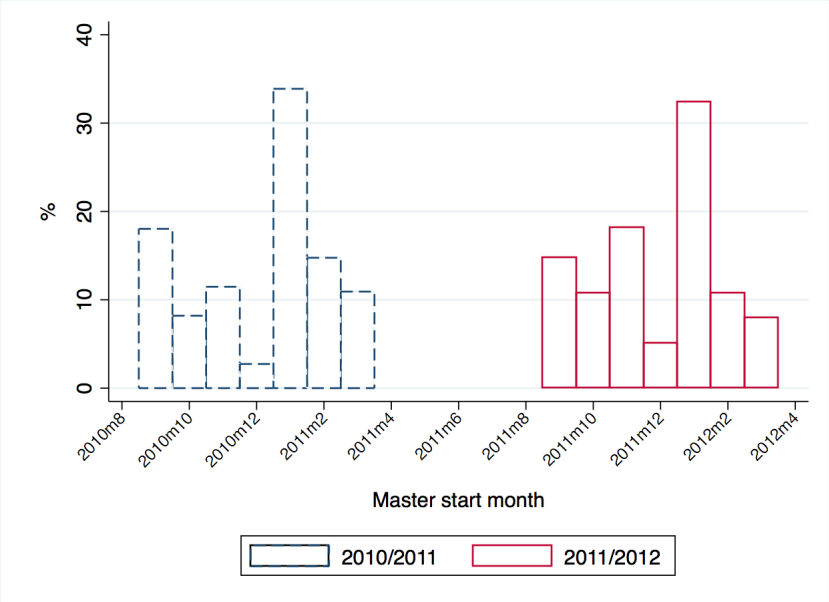
b. South/Central Italy



c. Abroad

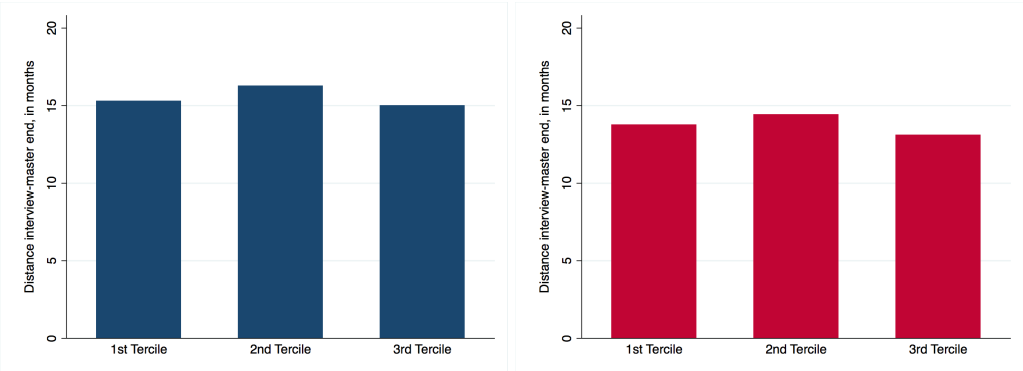
NOTES. This figure depicts the predicted probability of applying to a university in North, Central or South Italy and to a non-Italian university separately for the three cv-groups and for the two cohorts of applicants. The dots represent the point estimates, while the vertical lines the 90% confidence interval. The sample includes applicants who have started a master's degree within two years from their five-years degree. Students in social sciences, engineering and technology, arts and humanities, life sciences and medicine are included, while students in natural sciences are excluded. Regressions are estimated by OLS. Robust standard errors in parentheses. *** p< 0.01, ** p<0.05, and *p<0.10.

Figure 3.7: Distribution of master start date



NOTES. This figure depicts the distribution of the start date of master's degrees chosen by MAB applicants in 2010/11 and 2011/12. Students in social sciences, engineering and technology, arts and humanities, life sciences and medicine are included, while students in natural sciences are excluded.

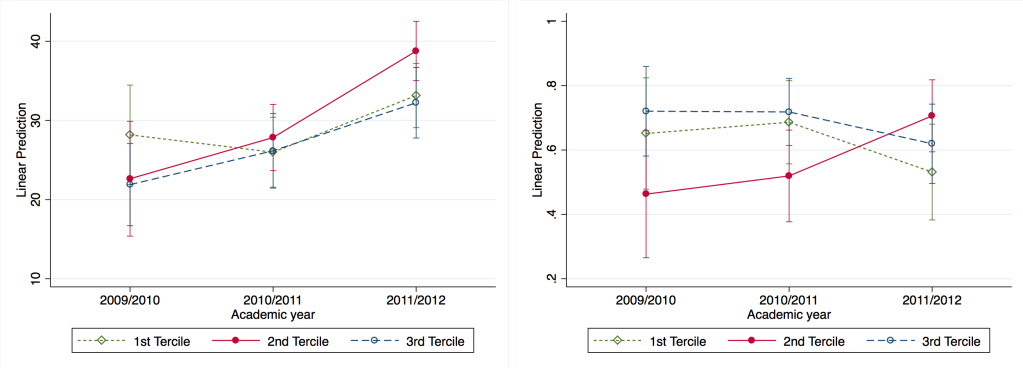
Figure 3.8: Interview date, by year and cv terciles



a. 2010/11 **b.** 2011/12

NOTES. This figure depicts the average time distance – in months – between the end date of master's degree chosen by MAB applicants in 2010/11 and 2011/12 and the first post-master interview, separately for the two cohorts and the cv groups. Students in social sciences, engineering and technology, arts and humanities, life sciences and medicine are included, while students in natural sciences are excluded.

Figure 3.9: Employment and university quality trends 2009/2010 - 2011/2012, by cv tercile



a. University quality **b.** Employment

NOTES. This figure depicts the linear prediction for the university score (Panel a) and the employment indicator (Panel b) for MAB applicants who attended a master's degree between the 2009/10 and the 2011/12 academic years. The university score ranges from 0 to 60, with universities listed behind the 300th position in the *THE-QS* rankings receiving a zero score and those listed at the 300th position and above receiving a score between 10 and 60. The employment indicator takes value one if the student is employed at the time of the *Almalaurea* interview. The average time distance between the end of the master's degree and the interview is 15 months. In both panels, dots and diamonds represent the predicted values from two linear regression (OLS) where the dependent variables are the university score and the employment indicator and the main independent variables are the interactions between two categorical variables: the one defining the three academic years and the one defining the cv terciles. Individual controls include subject, sex, age, family income, and an indicator for whether the final degree was achieved in a Sardinian university or not. The sample includes applicants who have started a master's degree within two years from their five-years degree. Students in social sciences, engineering and technology, arts and humanities, life sciences and medicine are included, while students in natural sciences are excluded. In Panel b., the sample is further restricted to applicants who completed their master's degree and for whom it is possible to determine the post-master employment situation.

Table 3.1: Grid for the determination of the cv score

Final degree grade	Points	**Duration of studies**	Points
100	0	On time	6
101	2	1 extra year	3
102	4	2 extra years	1
103	6	3 extra years or more	0
104	8	**Age at call**	Points
105	10	<25	4
106	12	25-27	3
107	14	28-30	2
108	16	31-33	1
109	18	≥ 34	0
110	19	**Further education**	Points
110 *magna cum laude*	20	Phd	4
Other experiences	Points	Specialisation course	4
Internships	0.10 per month	University master	3
(maximum 2 points)		Non-university master	2
Working experiences	0.10 per month	II level degree (extra)	2
(maximum 4 points)		I level master (in Italy)	1
		(maximum 4 points)	

NOTES. This Table reports the grid used to determine the cv score. The maximum number of points is 60. The components 'final degree grade' and 'duration of studies' refer to the five-year degree which is a compulsory requirement for being eligible for the MAB scholarship. The 'age at call' is computed as the age as of the call publication date.

Table 3.2: Applicants' Characteristics and Sample Selection

	2010/11		2011/2012		Difference	
	Mean	Sd	Mean	Sd	$Mean_{2011/12}$-$Mean_{2010/11}$	t
Cv-score	23.19	6.16	22.75	6.91	-0.44	(-0.63)
- Cv-score: final grade	17.12	4.76	16.65	5.25	-0.47	(-0.89)
- Cv-score: age at MAB call	2.51	0.82	2.66	0.87	0.15*	(1.69)
- Cv-score: duration of studies	2.50	2.29	2.54	2.44	0.04	(0.15)
- Cv-score: other studies/working exp.	1.20	1.55	1.11	1.56	-0.09	(-0.57)
Age at master start date	27.88	2.34	27.94	2.41	0.06	(0.22)
Female	0.60	0.49	0.60	0.49	0.00	(0.02)
Final grade (degree)	109.14	2.73	108.88	2.98	-0.26	(-0.87)
Age at degree	26.92	2.29	26.73	2.22	-0.19	(-0.78)
Degree completed on time	0.24	0.43	0.28	0.45	0.04	(0.94)
Graduate from a Sardinian university	0.70	0.46	0.66	0.48	-0.04	(-0.78)
Family income indicator (thousands of euros)	29.86	22.03	31.47	24.44	1.61	(0.66)
Family income indicator >80000 euros	0.08	0.27	0.10	0.30	0.03	(0.85)
Master's major subject: Arts & Humanities	0.12	0.33	0.09	0.29	-0.03	(-0.90)
Master's major subject: Life Sciences & Medicine	0.10	0.31	0.12	0.33	0.02	(0.63)
Master's major subject: Social Sciences	0.44	0.50	0.37	0.48	-0.07	(-1.41)
Master's major subject: Engineering & Technology	0.33	0.47	0.41	0.49	0.08	(1.59)
N. of applicants	183		176		359	
	Mean	Sd	Mean	Sd	$Mean_{2011/12}$-$Mean_{2010/11}$	t
Attended and completed the chosen master's degree	0.75	0.43	0.76	0.43	0.02	(0.35)
N. of applicants	137		133		270	
	Mean	Sd	Mean	Sd	$Mean_{2011/12}$-$Mean_{2010/11}$	t
Post-master survey	0.88	0.32	0.91	0.29	0.03	(0.71)
Distance interview-master end, in months	15.47	6.39	13.80	6.47	-1.67**	(-2.02)
N. of applicants	121		121		242	

NOTES. This table reports the baseline characteristics of the applicants for a MAB scholarship who attended a master's degree in the academic years 2010/11 and 2011/12. The sample average and the standard deviation are presented separately for the two cohorts of applicants. The last two columns report the mean difference between two cohorts and the t-statistic, respectively. In the upper panel the sample includes all applicants (359). The statistics about whether a post-master interview is available and the number of months between the interview date and the end of the master are computed for the sample of applicants who attended and completed the chosen master's degree (270). The sample includes applicants who have started a master's degree within two years from their five-years degree. Students in social sciences, engineering and technology, arts and humanities, life sciences and medicine are included, while students in natural sciences are excluded.
*** $p< 0.01$, ** $p<0.05$, and *$p<0.10$.

Table 3.3: Characteristics of the chosen master's degree

	2010/11		2011/2012		Difference	
	Mean	Sd	Mean	Sd	$Mean_{2011/12}$-$Mean_{2010/11}$	t
Program score	26.95	21.60	35.12	20.00	8.17***	(3.72)
Program score (as MAB 2011)	27.32	21.76	35.53	19.62	8.22***	(3.76)
Program score above 0/60	0.66	0.48	0.80	0.40	0.14***	(3.00)
Program score above 40/60	0.30	0.46	0.58	0.50	0.28***	(5.53)
Program score above 55/60	0.06	0.24	0.05	0.22	-0.01	(-0.37)
Master's tuition fees (thousands of euros)	8.06	9.87	6.62	5.62	-1.44*	(-1.69)
Master's duration (months)	12.18	2.54	12.80	2.66	0.62**	(2.24)
University Location: Italy	0.72	0.45	0.74	0.44	0.02	(0.48)
University Location: Europe	0.27	0.44	0.25	0.43	-0.02	(-0.38)
University Location: Outside Europe	0.02	0.13	0.01	0.11	-0.01	(-0.41)
Distance student's address-university (km)	781.60	795.11	786.48	938.64	4.88	(0.05)
Observations	183		176		359	

NOTES. This table reports the characteristics of the master's program chosen by applicants to the 2010 and 2011 editions of MAB. The sample average and the standard deviation are presented separately for the two cohorts of applicants. The last two columns report the mean difference between two cohorts and the t-statistic, respectively. The university score ranges from 0 to 60, with universities listed behind the 300th position in the *THE-QS* rankings receiving a zero score and those listed at the 300th position and above receiving a score between 10 and 60. The *university score (as MAB 2011)* is the *THE-QS* score computed following the 2010 rankings for both cohorts of applicants. The sample includes applicants who have started a master's degree within two years from their five-years degree. Students in social sciences, engineering and technology, arts and humanities, life sciences and medicine are included, while students in natural sciences are excluded.
*** $p< 0.01$, ** $p<0.05$, and *$p<0.10$.

Table 3.4: First stage

Dependent variable: University rank						
	University Score				Top 40% Uni.	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	TOBIT	OLS	LOGIT
MAB 2011/12	7.707***	8.326***	8.488***	9.176***	0.298***	1.424***
	(2.126)	(2.159)	(2.147)	(2.339)	(0.049)	(0.249)
Final grade (degree)		0.707*				
		(0.384)				
Degree completed on time		3.173				
		(2.772)				
Age at master start date		0.622				
		(0.519)				
Cv-score			0.181	0.197	0.006*	0.031
			(0.169)	(0.188)	(0.004)	(0.019)
Area of studies FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	No	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var.	30.959	30.882	30.882	30.882	0.436	0.436
Standard dev.	21.201	21.180	21.180	21.180	0.497	0.497
R^2	0.123	0.158	0.148		0.196	
Observations	359	358	358	358	358	358

NOTES. This table reports the coefficients of belonging to the 2011/12 cohort of MAB applicants on the rank of the chosen university. In Columns (1) to (4) the dependent variable is the score of the chosen university as computed by the Regional Administration following the *THE-QS* rankings. The university score ranges from 0 to 60, with universities listed behind the 300th position in the *THE-QS* rankings receiving a zero score and those listed at the 300th position and above receiving a score between 10 and 60. In Columns (5) and (6) the dependent variable is an indicator taking value one if the university is in the top 40% of the (area-specific) rankings, and zero otherwise. Individual controls include sex, age, family income, and an indicator for whether the final degree was achieved in a Sardinian university or not. The sample includes applicants to MAB who started a master's degree in 2010/11 and 2011/12 and within two years from their five-years degree. Students in social sciences, engineering and technology, arts and humanities, life sciences and medicine are included, while students in natural sciences are excluded. Regressions in Columns (1), (2), (3), and (5) are estimated by OLS. Regressions in Columns (4) and (6) are estimated following a TOBIT – left censored at a value of 10 – and a LOGIT specification, respectively. Column (6) reports the odds ratio. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$.

Table 3.5: Master completion

Dependent variable: Attended and completed the master's degree				
	ALL	T2 *vs* T3		T2 *vs* T1 and T3
	(1)	(2)	(3)	(4)
	OLS	OLS	OLS	OLS
MAB 2011/12	0.011	-0.012	0.005	0.004
	(0.046)	(0.060)	(0.061)	(0.056)
2nd Tercile		-0.189**	-0.183**	-0.063
		(0.075)	(0.078)	(0.072)
MAB 2011/12 × 2nd Tercile		0.057	0.050	0.054
		(0.101)	(0.100)	(0.098)
Area of studies FE	Yes	Yes	Yes	Yes
Individual controls	No	No	Yes	Yes
Mean Dep. Var.	0.756	0.815	0.814	0.756
Standard dev.	0.430	0.389	0.390	0.430
R^2	0.014	0.050	0.071	0.034
Observations	357	238	237	356

NOTES. This table reports the coefficients from the difference-in-differences Equation 3.1. In all Columns the dependent variable is an indicator taking value one if the student completes the master's degree, and zero if she does not. The indicator also takes value zero for students who, although having been admitted to a master's program, decide to not attend it, and when it is not possible to retrieve the information about a student's master completion. *MAB 2011/12* is a binary indicator taking value one if the MAB applicant attended a master's degree in the academic year 2011/12, while zero if she did so in 2010/11. 2nd *Tercile* is a binary indicator taking value one if the applicant belongs to the second tercile of the cv-score distribution, and zero otherwise. In Columns (2) and (3) the regressions are estimated on the sample of applicants belonging to the second and third tercile of the cv-score distribution only, while in Column (4) the sample also includes applicants in the bottom tercile. The sample includes applicants to MAB who started a master's degree in 2010/11 and 2011/12 and within two years from their five-years degree. Students in social sciences, engineering and technology, arts and humanities, life sciences and medicine are included, while students in natural sciences are excluded. Regressions in all Columns are estimated by OLS.
Robust standard errors in parentheses. *** p< 0.01, ** p<0.05, and *p<0.10.

Table 3.6: Cv-heterogeneity in university choice

Panel A: Tercile-specific effects						
	University score		Univ. score $\geq$40		Total Score	
	(1)		(2)		(3)	
MAB 2011/12 × 1st Tercile	6.835*		0.237***		5.355	
	(3.629)		(0.085)		(3.952)	
MAB 2011/12 × 2nd Tercile	10.627***		0.391***		10.269***	
	(3.413)		(0.083)		(3.354)	
MAB 2011/12 × 3rd Tercile	5.848		0.235***		5.787	
	(3.953)		(0.082)		(4.087)	
R^2	0.140		0.187		0.182	
Observations	358		358		358	
Panel B: Predicted Scores						
	2010/11	2011/12	2010/11	2011/12	2010/11	2011/12
1st Tercile	25.651	32.485	0.262	0.502	42.244	47.599
	(2.676)	(2.495)	(0.056)	(0.064)	(2.819)	(2.824)
2nd Tercile	27.698	38.326	0.249	0.632	52.194	62.464
	(2.539)	(2.325)	(0.055)	(0.058)	(2.524)	(2.260)
3rd Tercile	27.815	33.663	0.372	0.610	56.216	62.003
	(2.924)	(2.675)	(0.058)	(0.058)	(2.951)	(2.851)

NOTES. This table reports the coefficients of belonging to the 2011/12 cohort of MAB applicants on the rank of the chosen university, separately for the three cv groups. Panel A reports the coefficients of interaction between the indicator taking value one if the applicant attended the master's degree in the academic year 2011/12 – and zero if she did so in 2010/11 – and the cv-terciles dummies. The cv-terciles dummies are three binary indicators, each of them taking value one if an applicant belongs to a specific tercile, and zero otherwise. The dependent variable in Column (1) is the university score ranging from 0 to 60, with universities listed behind the 300th position in the *THE-QS* rankings receiving a zero score and those listed at the 300th position and above receiving a score between 10 and 60. The dependent variable in Column (2) is an indicator taking value one if the university is in the top 40% of the (area-specific) rankings, and zero otherwise. The dependent variable in Column (3) is the total score, defined as the sum of the cv and the university score and thus ranging from 0 to 100. Panel B reports the predicted university and total score, and the predicted probability of choosing a university listed in the top 40% of the rankings separately for the three cv-groups and for the two cohorts of applicants. The sample includes applicants to MAB who started a master's degree in 2010/11 and 2011/12 and within two years from their five-years degree. Students in social sciences, engineering and technology, arts and humanities, life sciences and medicine are included, while students in natural sciences are excluded. Regressions are estimated by OLS. Robust standard errors in parentheses. *** p< 0.01, ** p<0.05, and *p<0.10.

Table 3.7: Main result

Dependent variable: Employed				
	ALL	T2 *vs* T3		T2 *vs* T1 and T3
	(1)	(2)	(3)	(4)
	OLS	OLS	OLS	OLS
MAB 2011/12	-0.019	-0.102	-0.112	-0.141*
	(0.061)	(0.098)	(0.100)	(0.077)
2nd Tercile		-0.203*	-0.221**	-0.196**
		(0.106)	(0.105)	(0.097)
MAB 2011/12 × 2nd Tercile		0.287*	0.294**	0.319**
		(0.146)	(0.145)	(0.131)
Area of studies FE	Yes	Yes	Yes	Yes
Individual controls	No	No	Yes	Yes
Mean Dep. Var.	0.640	0.651	0.651	0.640
Standard dev.	0.481	0.478	0.478	0.481
R^2	0.033	0.060	0.086	0.076
Observations	242	175	175	242

NOTES. This table reports the coefficients from the difference-in-differences Equation 3.1. In all Columns the dependent variable is an indicator taking value one if the student is employed at the time of the *Almalaurea* interview. The average time distance between the end of the master's degree and the interview is 15 months. *MAB 2011/12* is a binary indicator taking value one if the MAB applicant attended a master's degree in the academic year 2011/12, while zero if she did so in 2010/11. *2nd Tercile* is a binary indicator taking value one if the applicant belongs to the second tercile of the cv-score distribution, and zero otherwise. In Columns (2) and (3) the regressions are estimated on the sample of applicants belonging to the second and third tercile of the cv-score distribution only, while in Column (4) the sample also includes applicants in the bottom tercile. The sample includes applicants to MAB who started a master's degree in 2010/11 and 2011/12 and within two years from their five-years degree. It is further limited to applicants who completed their master's degree and for whom at least one post-master interview is available. Students in social sciences, engineering and technology, arts and humanities, life sciences and medicine are included, while students in natural sciences are excluded. Regressions in all Columns are estimated by OLS.
Robust standard errors in parentheses. *** p< 0.01, ** p<0.05, and *p<0.10.

Table 3.8: Triple-differences estimates

Dependent variable: Post-master employment situation			
	ALL	T2 *vs* T3	T2 *vs* T1 and T3
	(1)	(2)	(3)
	OLS	OLS	OLS
MAB Applicant	-0.178***	-0.168**	-0.123**
	(0.050)	(0.069)	(0.058)
Master in 2011/12	-0.034*	-0.035	-0.027
	(0.020)	(0.032)	(0.025)
MAB Applicant × Master in 2011/12	-0.014	-0.093	-0.101
	(0.067)	(0.100)	(0.084)
2nd Tercile		0.002	0.024
		(0.035)	(0.032)
MAB Applicant × 2nd Tercile		-0.170	-0.187*
		(0.104)	(0.098)
Master in 2011/12 × 2nd Tercile		-0.022	-0.027
		(0.053)	(0.049)
MAB Applicant × Master in 2011/12 × 2nd Tercile		0.258*	0.267*
		(0.155)	(0.144)
Area of studies FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes
Mean Dep. Var.	0.784	0.789	0.787
Standard dev.	0.411	0.408	0.409
R^2	0.099	0.136	0.122
Observations	1763	1063	1489

NOTES. This table reports the coefficients from the triple-differences Equation 3.2. In all Columns the dependent variable is an indicator taking value one if the student is employed at the time of the *Almalaurea* interview. *MAB applicant* is a binary indicator taking value one if the student is a Sardinian MAB applicant, and zero when she is a non-Sardinian postgraduate student. *Master in 2011/12* is a binary indicator taking value one if the student attended a master's degree in the academic year 2011/12, while zero if she did so in 2010/11. *2nd Tercile* is a binary indicator taking value one if the applicant belongs to the second tercile of the cv-score distribution, and zero otherwise. The terciles are computed following the cv-score distribution of MAB applicants in 2010/11 and 2011/12. In Columns (1) and (2) the regressions are estimated on the sample of applicants belonging to the second and third tercile of the cv-score distribution only, while in Column (3) the sample also includes applicants in the bottom tercile. The sample includes postgraduate students who started and completed a master's degree in 2010/11 and 2011/12 and within two years from their five-years degree. Students in social sciences, engineering and technology, arts and humanities, life sciences and medicine are included, while students in natural sciences are excluded. Besides Sardininan MAB applicants, only students from regions for which *Almalaurea* provides a coverage of at least 80% are considered. These regions are Trentino Alto-Adige, Veneto, Friuli-Venezia Giulia, Liguria, Emilia-Romagna, and Umbria. Regressions in all Columns are estimated by OLS.
Robust standard errors in parentheses. *** p< 0.01, ** p<0.05, and *p<0.10.

Bibliography

Agarwal, S., Chomsisengphet, S., Mahoney, N., and Stroebel, J. (2015). Regulating Consumer Financial Products: Evidence from Credit Cards. *The Quarterly Journal of Economics*, 130(1):111–164.

Almenberg, J. and Dreber, A. (2015). Gender, Stock Market Participation and Financial Literacy. *Economics Letters*, 137(C):140–142.

Andrews, R. J., Li, J., and Lovenheim, M. F. (2016). Quantile treatment effects of college quality on earnings. *Journal of Human Resources*, 51(1):200–238.

Ashraf, N., Bandiera, O., and Lee, S. (2018). Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services. Working paper.

Atkinson, A. and Messy, F.-A. (2012). Measuring Financial Literacy: Results of the OECD / International Network on Financial Education (INFE) Pilot Study. OECD Working Paper on Finance, Insurance and Private Pensions No. 15, OECD.

Bagues, M., Sylos-Labini, M., and Zinovyeva, N. (2017). Does the Gender Composition of Scientific Committees Matter? *American Economic Review*, 107(4):1207–1238.

Baicker, K. and Jacobson, M. (2007). Finders Keepers: Forfeiture Laws, Policing Incentives, and Local Budgets. *Journal of Public Economics*, 91(11):2113–2136.

Behrman, J. R., Roswnzweig, M. R., and Taubman, P. (1996). College choice and wages: Estimates using data on female twins. *The Review of Economics and Statistics*, 78(4):672–685.

Bernheim, B. D. (1995). Do Households Appreciate their Financial Vulnerabilities? An Analysis of Actions, Perceptions, and Public Policy. *Tax Policy and Economic Growth*, pages 1–30.

Bernheim, B. D., Garrett, D. M., and Maki, D. M. (2001). Education and Saving:: The Long-term Effects of High School Financial Curriculum Mandates. *Journal of Public Economics*, 80(3):435–465.

Black, D. A. and Smith, J. A. (2006). Estimating the returns to college quality with multiple proxies for quality. *Journal of Labor Economics*, 24(3).

Bleemer, Z. and Zafar, B. (2018). Intended college attendance: Evidence from an experiment on college returns and costs. *Journal of Public Economics*, 157:184–211.

Borjas, G. J. and Doran, K. (2015). Which Peers Matter? The Relative Impacts of Collaborators, Colleagues, and Competitors. *The Review of Economics and Statistics*, 97(5):1104–1117.

Brown, M., Grigsby, J., van der Klaauw, W., Wen, J., and Zafar, B. (2016). Financial Education and the Debt Behavior of the Young. *Review of Financial Studies*, 29(9):2490–2522.

Brugiavini, A., Cavapozzi, D., Padula, M., and Pettinicchi, Y. (2015). Financial Education, Literacy and Investment Attitudes. SAFE Working Paper No. 86, Research Center SAFE - Sustainable Architecture for Finance in Europe, Goethe University Frankfurt.

Brunello, G. and Cappellari, L. (2008). The labour market effects of Alma Mater: Evidence from Italy. *Economics of Education Review*, 27(5):564–574.

Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2017). Rdrobust: Software for Regression-Discontinuity Designs. *Stata Journal*, 17(2):372–404.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, 82(6):2295–2326.

Calvet, L. E., Campbell, J. Y., and Sodini, P. (2007). Down or Out: Assessing the Welfare Costs of Household Investment Mistakes. *Journal of Political Economy*, 115(5):707–747.

Calvet, L. E., Campbell, J. Y., and Sodini, P. (2009). Measuring the Financial Sophistication of Households. *American Economic Review*, 99(2):393–398.

Campbell, D. (2008). Nonfinancial Performance Measures and Promotion-Based Incentives. *Journal of Accounting Research*, 46(2):297–332.

Campbell, J. Y. (2006). Household Finance. *Journal of Finance*, 61(4):1553–1604.

Canaan, S. and Mouganie, P. (2018). Returns to education quality for low-skilled students: Evidence from a discontinuity. *Journal of Labor Economics*, 36(2):395–436.

Carpena, F., Cole, S., Shapiro, J., and Zia, B. (2018). The ABCs of Financial Education: Experimental Evidence on Attitudes, Behavior, and Cognitive Biases. *Management Science,, forthcoming*.

Cattaneo, M. D., Jansson, M., and Ma, X. (2017a). Rddensity: Manipulation Testing Based on Density Discontinuity. *Stata Journal*, page *forthcoming*.

Cattaneo, M. D., Jansson, M., and Ma, X. (2017b). Simple Local Polynomial Density Estimators. Working paper, University of Michigan.

Choi, J. J., Laibson, D., and Madrian, B. C. (2011). $100 bills on the sidewalk: Suboptimal investment in 401 (k) plans. *Review of Economics and Statistics*, 93(3):748–763.

Christelis, D., Georgarakos, D., and Haliassos, M. (2011). Stockholding: Participation, Location, and Spillovers. *Journal of Banking & Finance*, 35(8):1918 – 1930.

Christelis, D., Jappelli, T., and Padula, M. (2010). Cognitive Abilities and Portfolio Choice. *European Economic Review*, 54(1):18–38.

Christiansen, C., Joensen, J. S., and Rangvid, J. (2008). Are Economists More Likely to Hold Stocks? *Review of Finance*, 12(3):465–496.

Cohodes, S. R. and Goodman, J. S. (2014). Merit aid, college quality, and college completion: Massachusetts' adams scholarship as an in-kind subsidy. *American Economic Journal: Applied Economics*, 6(4):251–85.

Cole, S., Paulson, A., and Shastry, G. (2014). Smart Money? The Effect of Education on Financial Outcomes. *Review of Financial Studies*, 27(7):2022–2051.

Collins, J. M. (2013). The Impacts of Mandatory Financial Education: Evidence from a Randomized Field Study. *Journal of Economic Behavior & Organization*, 95(C):146–158.

Dal Bó, E., Finan, F., and Rossi, M. (2013). Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service. *The Quarterly Journal of Economics*, 128(3):1169–1218.

Dale, S. B. and Krueger, A. B. (2002). Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables. *The Quarterly Journal of Economics*, (November).

Dale, S. B. and Krueger, A. B. (2014). Estimating the effects of college characteristics over the career using administrative earnings data. *Journal of human resources*, 49(2):323–358.

De Paola, M., Ponzo, M., and Scoppa, V. (2017). Gender Differences in the Propensity to Apply for Promotion: Evidence from the Italian Scientific Qualification. *Oxford Economic Papers*, 69(4):986–1009.

De Paola, M. and Scoppa, V. (2015). Gender Discrimination and Evaluators' Gender: Evidence from Italian Academia. *Economica*, 82(325):162–188.

Deuflhard, F., Georgarakos, D., and Inderst, R. (2018). Financial Literacy and Savings Account Returns. *Journal of the European Economic Association,*, *forthcoming*.

Dimmock, S. G., Kouwenberg, R., Mitchell, O. S., and Peijnenburg, K. (2016). Ambiguity Aversion and Household Portfolio Choice Puzzles: Empirical Evidence. *Journal of Financial Economics*, 119(3):559 – 577.

Drexler, A., Fischer, G., and Schoar, A. (2014). Keeping It Simple: Financial Literacy and Rules of Thumb. *American Economic Journal: Applied Economics*, 6(2):1–31.

Duflo, E., Hanna, R., and Ryan, S. (2012). Incentives Work: Getting Teachers to Come to School. *American Economic Review*, 102(4):1241–78.

Duflo, E. and Saez, E. (2003). The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment. *The Quarterly Journal of Economics*, 118(3):815–842.

Easley, D. and O'Hara, M. (2010). Microstructure and Ambiguity. *Journal of Finance*, 65(5):1817–1846.

Faria, J. R. and McAdam, P. (2014). Does Tenure Make Researchers Less Productive? The Case of the "Specialist". School of Economics Discussion Papers 0514, School of Economics, University of Surrey.

Fernandes, D., Lynch, J. G., and Netemeyer, R. G. (2014). Financial Literacy, Financial Education, and Downstream Financial Behaviors. *Management Science*, 60(8):1861–1883.

Finan, F., Olken, B. A., and Pande, R. (2017). The Personnel Economics of the Developing State. *Handbook of Economic Field Experiments*, 2:467–514.

Fort, M., Manaresi, F., and Trucchi, S. (2016). Adult Financial Literacy and Households' Financial Assets: the Role of Bank Information Policies. *Economic Policy*, 31(88):743–782.

Friebel, G., Heinz, M., Krueger, M., and Zubanov, N. (2017). Team Incentives and Performance: Evidence from a Retail Chain. *American Economic Review*, 107(8):2168–2203.

Gaudecker, H.-M. v. (2015). How Does Household Portfolio Diversification Vary with Financial Literacy and Financial Advice? *The Journal of Finance*, 70(2):489–507.

Gaynor, M., Rebitzer, J., and Taylor, L. J. (2004). Physician Incentives in Health Maintenance Organizations. *Journal of Political Economy*, 112(4):915–931.

Gertler, P. and Vermeersch, C. (2013). Using Performance Incentives to Improve Medical Care Productivity and Health Outcomes. NBER Working Papers 19046, National Bureau of Economic Research, Inc.

Glewwe, P., Ilias, N., and Kremer, M. (2010). Teacher incentives. *American Economic Journal: Applied Economics*, 2(3):205–27.

Goodman, S. (2016). Learning from the test: Raising selective college enrollment by providing information. *Review of Economics and Statistics*, 98(4):671–684.

Guiso, L. and Jappelli, T. (2009). Financial Literacy and Portfolio Diversification. CSEF Working Paper No. 212, Centre for Studies in Economics and Finance, University of Naples.

Haeck, C. and Verboven, F. (2012). The Internal Economics of a University: Evidence from Personnel Data. *Journal of Labor Economics*, 30(3):591 – 626.

Haliassos, M., Jansson, T., and Karabulut, Y. (2017). Financial Literacy Externalities. CEPR Discussion Paper No. DP12100, Centre for Economic Policy Research.

Hastings, J. S., Madrian, B. C., and Skimmyhorn, W. L. (2013a). Financial Literacy, Financial Education, and Economic Outcomes. *Annual Review of Economics*, 5(1):347–373.

Hastings, J. S., Neilson, C. A., and Zimmerman, S. D. (2013b). Are some degrees worth more than others? evidence from college admission cutoffs in chile. Technical report, National Bureau of Economic Research.

Hershbein, B. J. (2013). Worker signals among new college graduates: The role of selectivity and GPA. Technical report.

Hoekstra, M. (2009). The effect of attending the flagship state university on earnings: A discontinuity-based approach. *The Review of Economics and Statistics*, 91(November):717–724.

Holt, C. A. and Laury, S. K. (2002). Risk Aversion and Incentive Effects. *American Economic Review*, 92(5):1644–1655.

Izhakian, Y. and Yermack, D. (2017). Risk, Ambiguity, and the Exercise of Employee Stock Options. *Journal of Financial Economics*, 124(1):65–85.

Jappelli, T. and Padula, M. (2013). Investment in Financial Literacy and Saving Decisions. *Journal of Banking & Finance*, 37(8):2779–2792.

Karachiwalla, N. and Park, A. (2017). Promotion Incentives in the Public Sector: Evidence from Chinese Schools. *Journal of Public Economics*, 146(C):109–128.

Kimball, M. and Shumway, T. (2006). Investor Sophistication, and the Participation, Home Bias, Diversification, and Employer Stock Puzzles. Mimeo, University of Michigan.

Klapper, L., Lusardi, A., and Panos, G. A. (2013). Financial Literacy and its Consequences: Evidence from Russia During the Financial Crisis. *Journal of Banking & Finance*, 37(10):3904–3923.

Klapper, L., Lusardi, A., and Van Oudheusden, P. (2015). Financial Literacy Around the World. Global financial literacy excellence center report, George Washington University.

Klibanoff, P., Marinacci, M., and Mukerji, S. (2005). A Smooth Model of Decision Making under Ambiguity. *Econometrica*, 73(6):1849–1892.

Kwon, I. (2006). Incentives, Wages, and Promotions: Theory and Evidence. *The RAND Journal of Economics*, 37(1):100–120.

Lang, K. and Siniver, E. (2011). Why is an elite undergraduate education valuable? Evidence from Israel. *Labour Economics*, 18(6):767–777.

Lavy, V. (2002). Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement. *Journal of Political Economy*, 110(6):1286–1317.

Lazear, E. P. (2000). Performance Pay and Productivity. *American Economic Review*, 90(5):1346–1361.

Lazear, E. P. and Gibbs, M. (2014). *Personnel Economics in Practice*. John Wiley & Sons.

Lazear, E. P. and Oyer, P. (2013). Personnel economics. In Gibbons, R. and Roberts, J., editors, *The Handbook of Organizational Economics*. Princeton University Press.

Lindahl, L. and Regner, H. (2005). College Choice and Subsequent Earnings: Results Using Swedish Sibling Data. *Scandinavian Journal of Economics*, 107(3):437–457.

Lusardi, A. and de Bassa Scheresberg, C. (2013). Financial Literacy and High-Cost Borrowing in the United States. NBER Working Paper No. 18969, National Bureau of Economic Research.

Lusardi, A., Michaud, P.-C., and Mitchell, O. S. (2017). Optimal Financial Knowledge and Wealth Inequality. *Journal of Political Economy*, 125(2):431–477.

Lusardi, A. and Mitchell, O. S. (2007). Financial Literacy and Retirement Preparedness: Evidence and Implications for Financial Education. *Business Economics*, 42(1):35–44.

Lusardi, A. and Mitchell, O. S. (2010). How Ordinary Consumers Make Complex Economic Decisions: Financial Literacy and Retirement Readiness. CFS Working Paper No. 2010/11, Center for Financial Studies (CFS).

Lusardi, A. and Mitchell, O. S. (2011). Financial Literacy and Planning: Implications for Retirement Wellbeing. NBER Working Paper No. 17078, National Bureau of Economic Research.

Lusardi, A. and Mitchell, O. S. (2014). The Economic Importance of Financial Literacy: Theory and Evidence. *Journal of Economic Literature*, 52(1):5–44.

Lusardi, A. and Tufano, P. (2015). Debt Literacy, Financial Experiences, and Overindebtedness. *Journal of Pension Economics and Finance*, 14(04):332–368.

Lührmann, M., Serra-Garcia, M., and Winter, J. (2015). Teaching Teenagers in Finance: Does it Work? *Journal of Banking & Finance*, 54(C):160–174.

Maccheroni, F., Marinacci, M., and Ruffino, D. (2013). Alpha as Ambiguity: Robust Mean-Variance Portfolio Analysis. *Econometrica*, 81(3):1075–1113.

MacLeod, W. B., Riehl, E., Saavedra, J. E., and Urquiola, M. (2017). The big sort: College reputation and labor market outcomes. *American Economic Journal: Applied Economics*, 9(3):223–61.

McCrary, J. (2008). Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics*, 142(2):698–714.

Moore, D. L. (2003). *Survey of Financial Literacy in Washington State: Knowledge, Behavior, Attitudes, and Experiences*. Washington State Department of Financial Institutions.

Moser, P., Voena, A., and Waldinger, F. (2014). German Jewish Èmigrès and US Invention. *American Economic Review*, 104(10):3222–3255.

Muralidharan, K. and Sundararaman, V. (2011). Teacher Performance Pay: Experimental Evidence from India. *Journal of Political Economy*, 119(1):39 – 77.

OECD (2017). *Government at a Glance 2017*. OECD Publishing.

Olken, B. A., Onishi, J., and Wong, S. (2014). Should Aid Reward Performance? Evidence from a Field Experiment on Health and Education in Indonesia. *American Economic Journal: Applied Economics*, 6(4):1–34.

Oosterbeek, H. and Webbink, D. (2011). Does Studying Abroad Induce a Brain Drain? *Economica*, 78(310):347–366.

Pallais, A. (2014). Small differences that matter: Mistakes in applying to college. *Journal of Labor Economics*, 33(2):38.

Papay, J. P., Murnane, R. J., and Willett, J. B. (2011). Extending the Regression-Discontinuity Approach to Multiple Assignment Variables. *Journal of Econometrics*, 161(2):203–207.

Papay, J. P., Murnane, R. J., and Willett, J. B. (2014). High-School Exit Examinations and the Schooling Decisions of Teenagers: Evidence From Regression-Discontinuity Approaches. *Journal of Research on Educational Effectiveness*, 7(1):1–27.

Parey, M. and Waldinger, F. (2011). Studying Abroad and the Effect on International Labour Market Mobility: Evidence from the Introduction of ERASMUS*. *The Economic Journal*, 121:194–222.

Rodríguez, J., Urzúa, S., and Reyes, L. (2016). Heterogeneous economic returns to post-secondary degrees: Evidence from chile. *Journal of Human Resources*, 51(2):416–460.

Saavedra, J. E. (2009). The learning and early labor market effects of college quality: A regression discontinuity analysis.

Sayinzoga, A., Bulte, E. H., and Lensink, R. (2016). Financial Literacy and Financial Behaviour: Experimental Evidence from Rural Rwanda. *The Economic Journal*, 126(594):1571–1599.

Sekita, S. (2011). Financial Literacy and Retirement Planning in Japan. *Journal of Pension Economics and Finance*, 10(04):637–656.

Shearer, B. (2004). Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment. *The Review of Economic Studies*, 71(2):513–534.

Song, C. (2015). Financial Illiteracy and Pension Contributions: A Field Experiment on Compound Interest in China. Mimeo, University of California, Berkeley.

van Rooij, M., Lusardi, A., and Alessie, R. (2011). Financial Literacy and Stock Market Participation. *Journal of Financial Economics*, 101(2):449–472.

Waldinger, F. (2012). Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany. *Review of Economic Studies*, 79(2):838–861.

Waldinger, F. (2016). Bombs, Brains, and Science: The Role of Human and Physical Capital for the Creation of Scientific Knowledge. *The Review of Economics and Statistics*, 98(5):811–831.

Zimmerman, S. D. (2014). The returns to college admission for academically marginal students. *Journal of Labor Economics*, 32(4):711–754.

Zinovyeva, N. and Bagues, M. (2015). The Role of Connections in Academic Promotions. *American Economic Journal: Applied Economics*, 7(2):264–292.