Essays in Labor Economics

Brandimarti, Eleonora

This publication URL: https://archive-ouverte.unige.ch/unige:170170
Publication DOI: 10.13097/archive-ouverte/unige:170170

Essays in Labor Economics

Dissertations d'Économie du Travail

THESE
soumise à la
Geneva School of Economics and Management,
Université de Genève, Suisse,

Par

Eleonora BRANDIMARTI

Sous la direction de
Prof. Michele PELLIZZARI, co-directeur de thèse
et
Prof. Giacomo DE GIORGI, co-directeur de thèse

a répondu aux conditions requises pour obtenir le

Docteur en économie et management
mention *Economie*

Membres du jury
Prof. Aleksey TETENOV, président du jury, Université de Genève
Prof. Giacomo DE GIORGI, co-directeur de thèse, Université de Genève
Prof. Michele PELLIZZARI, co-directeur de thèse, Université de Genève
Prof. Edwin LEUVEN, Oslo University, Norway

Thèse no. 121
Genève, *Juin 2023*

La Faculté d'économie et de management, sur préavis du jury, a autorisé l'impression de la présente thèse, sans entendre, par-là, émettre aucune opinion sur les propositions qui s'y trouvent énoncées et qui n'engagent que la responsabilité de leur auteur.

Genève, le 13 juin 2023

Doyen
Markus MENZ

Acknowledgements

Undertaking this PhD has been an all-encompassing and transformative experience for which I am indebted to my family, friends, and colleagues.

First, I would like to express my sincere gratitude to my advisors Prof. Michele Pellizzari and Prof. Giacomo De Giorgi for their excellent supervision. Their guidance, mentorship, support, and sharing of their immense knowledge has been crucial to the development of my research and to my growth as a researcher. I am particularly grateful for their kind assistance throughout the pandemic and the job market; without them, none of this would have been achievable. I would also like to thank the other members of my committee, Prof. Aleksey Tetenov and Prof. Edwin Leuven, for their thoughtful insights that widened my perspective.

I am grateful to many other faculty members and professors with whom I had the opportunity to interact. My gratitude extends to Prof. Peter Arcidiacono, whose insight and feedback was crucial to advance in Chapter 2, Prof. Jérémy Laurent-Lucchetti, whose exciting research agenda has been a constant source of inspiration, Prof. Aleksey Tetenov, who has doctored all my research ideas, and Prof. Julien Daubanes and Prof. Frédéric Robert-Nicoud, who offered invaluable support throughout the job market. I am thankful to Federica Sbergami, who taught me how to teach and always offered kind support, and to my coauthors Gaetano Basso and Prof. Giovanni Pica for their insight and help in completing Chapter 1.

Countless conversations with my friends and colleagues have inspired me over the years and made my time as a PhD candidate so enjoyable. In particular, I thank Rasha Shakra, Federica Braccioli, Avichal Mahajan, Flavia Cifarelli, Pietro Campa, Davide Pietrobon, Tammaro Terracciano, Kyungbo Han, and the Rare Voices in Economics family. Silvia Ghiselli, Davide Cristofori, and Claudia Girotti offered precious support in accessing the AlmaLaurea resources and carrying out many research projects. Friends from outside the PhD have often helped me take a much-needed break from research; I am grateful to Fulvia Serra, Giulia Simoni, and Giorgia Gazza for being my greatest cheerleaders.

I thank, from the bottom of my heart, my beloved parents and siblings for their unconditional and everlasting support. My mother, Marianna, for always asking the hard questions, my father, Renato, for his contagious curiosity in searching for answers, and Federico, Melaku, and Matilde for challenging me to think outside of the box.

Finally, I am forever indebted to Stefano, who by now probably knows more about selection problems and the academic job market than he would have ever wished. I could never have completed this adventure without his unconditional love, support, and patience.

Abstract

This thesis focuses on how selection into opportunities and self-selection affect labor market outcomes, particularly focusing on early career and political outcomes. While work and voting decisions are motivated by different sets of determinants, they are strongly affected by institutions that are fundamentally present in modern and democratic economies. Here, I present three instances in which individual decisions – affected by preferences and opportunities – determine collective outcomes.

In the first part of the thesis, we investigate a mechanism to rationalize the empirical evidence of the market failure of occupational licensing. Entry in many occupations is regulated to screen out the least able producers but the available evidence suggests that this objective is rarely achieved. Using microdata covering the universe of Italian law school graduates (2007-2013), we show that this result is due to the strong intergenerational transmission of regulated professions. We find that having relatives already active in the profession substantially increases the probability of passing the entry exam (and earnings), especially so for those who performed poorly in law school. We do not find evidence of intergenerational transmission of occupation-specific human capital. Counterfactual simulations show that positive selection emerges if family connections are assumed away.

In the second part of the thesis, I investigate how knowledge taught at university beyond degrees affects the labor market outcomes of graduates. Using novel data covering the universe of Italian graduates, I find that returns to combinations of bachelor's and master's degrees vary substantially even for combinations with the same undergraduate degree, suggesting that both types of programs require consideration. Multidisciplinary university careers relate positively to economic outcomes, while combinations in the same field perform worse. Quantitative courses alone do not explain higher returns.

In the third part of the thesis, we investigate how uncertainty in access to credit affects political outcomes in the US. There is a tight connection between credit access and voting. We show that uncertainty in access to credit pushes voters toward more conservative candidates in US elections. Using a 1% sample of the US population with valid credit reports, we relate access to credit to voting outcomes in all county-by-congressional districts over the period 2004-2016. Specifically, we construct exogenous measures of uncertainty to credit access, i.e. credit score values around which individual total credit amount jumps the most (e.g. around which uncertainty on access to credit is the highest). We then show that a 10pp increase in the share of marginal voters located just around these thresholds increases Republican votes by 2.7pp and reduces that of Democrats by 2.6pp. Furthermore, winning candidates in more uncertain constituencies tend to follow a more conservative rhetoric.

Resumé

Cette thèse se concentre sur la manière dont les choix et l'autosélection affectent les résultats sur le marché du travail, en se concentrant particulièrement sur les débuts de carrière et les résultats politiques. Bien que les décisions de travail et de vote soient motivées par des déterminants différents, elles sont fortement influencées par les institutions qui sont fondamentalement présentes dans les économies modernes et démocratiques. Je présente ici trois exemples dans lesquels les décisions individuelles - influencées par les préférences et les opportunités - déterminent les résultats collectifs.

Dans la première partie de la thèse, nous étudions un mécanisme permettant de rationaliser les preuves empiriques de l'échec du marché des licences professionnelles. L'accès à de nombreuses professions est réglementé afin d'écarter les producteurs les moins compétents, mais les données disponibles suggèrent que cet objectif est rarement atteint. En utilisant des microdonnées couvrant l'univers des diplômés des facultés de droit italiennes (2007-2013), nous montrons que ce résultat est dû à la forte transmission intergénérationnelle des professions réglementées. Nous constatons que le fait d'avoir des parents déjà actifs dans la profession augmente considérablement la probabilité de réussir l'examen d'entrée (et les revenus), en particulier pour ceux qui ont eu de mauvais résultats à la faculté de droit. Nous ne trouvons pas de preuve de la transmission intergénérationnelle du capital humain spécifique à la profession. Des simulations contrefactuelles montrent qu'une sélection positive émerge si l'on ne tient pas compte des liens familiaux.

Dans la deuxième partie de la thèse, j'étudie comment les connaissances enseignées à l'université affectent les résultats des diplômés sur le marché du travail au-delà des diplômes. En utilisant de nouvelles données couvrant l'univers des diplômés italiens, je constate que les rendements des combinaisons de diplômes de licence et de master varient considérablement, même pour les combinaisons avec le même diplôme de premier cycle, ce qui suggère que les deux types de programmes doivent être pris en considération. Les carrières universitaires pluridisciplinaires sont liées positivement aux résultats économiques, tandis que les combinaisons dans le même domaine sont moins performantes. Les cours quantitatifs n'expliquent pas à eux seuls les rendements plus élevés.

Dans la troisième partie de la thèse, nous étudions comment l'incertitude de l'accès au crédit affecte les résultats politiques aux États-Unis. Il existe un lien étroit entre l'accès au crédit et le vote. Nous montrons que l'incertitude de l'accès au crédit pousse les électeurs vers des candidats plus conservateurs lors des élections américaines. En utilisant un échantillon de 1% de la population américaine avec des rapports de crédit valides, nous établissons un lien entre l'accès au crédit et les résultats du vote dans toutes les circonscriptions électorales par comté sur la période 2004-2016. Plus précisément, nous construisons des mesures exogènes de l'incertitude de l'accès au crédit, c'est-à-dire des valeurs de score de crédit autour desquelles le montant total du crédit individuel augmente le plus (c'est-à-dire autour desquelles l'incertitude de l'accès au crédit est la plus élevée). Nous montrons ensuite qu'une augmentation de 10 points de pourcentage de la part des électeurs marginaux situés juste autour de ces seuils augmente les votes républicains de 2,7 points de pourcentage et réduit ceux des démocrates de 2,6 points de pourcentage. En outre, les candidats victorieux dans les circonscriptions plus incertaines ont tendance à suivre une rhétorique plus conservatrice.

Contents

List of Figures

List of Tables

CHAPTER 1

Quality and Selection in Regulated Professions

Joint with Gaetano Basso, Michele Pellizzari, and Giovanni Pica

Entry in many occupations is regulated to screen out the least able producers but the available evidence suggests that this objective is rarely achieved. Using microdata covering the universe of Italian law school graduates (2007-2013), we show that this result is due to the strong intergenerational transmission of regulated professions. We find that having relatives already active in the profession substantially increases the probability of passing the entry exam (and earnings), especially so for those who performed poorly in law school. We do not find evidence of intergenerational transmission of occupation-specific human capital. Counterfactual simulations show that positive selection emerges if family connections are assumed away.[1]

1.1 Introduction

Entry in many occupations is regulated with the objective to protect consumers by selecting only the most able producers into the market (Friedman and Kuznets, 1945; Kleiner and Krueger, 2013; Bryson and Kleiner, 2010; Kleiner,

[1]We are grateful to Silvia Ghiselli and the AlmaLaurea team for granting us access to their databases, to Marco Jazzetta for helping us with data collection and to Paolo Buonanno&Mario Pagliero and Michele Raitano&Francesco Vona for sharing some of their data and results with us. We would like to thank participants in the Bank of Italy Workshop "The Economics of Occupational Licensing", the Genoa Spring Workshop in "Labour Market Institutions", the Alp-Pop 2020 and the ASSA 2021 meetings for useful comments and suggestions. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Bank of Italy or of the Eurosystem. Any errors or omissions are the sole responsibility of the authors.

2000). However, the available empirical evidence suggests that in most cases occupational regulation fails to achieve such a goal (Anderson et al., 2020; Kleiner, 2017; Bryson and Kleiner, 2019). The robustness of this finding across different professions and institutional contexts is indeed surprising, as occupational regulations are explicitly designed to produce positive selection, and solid theoretical considerations suggest they should work (Leland, 1979; Maurizi, 1974; Shapiro, 1986; Stigler, 1971).[2] Nonetheless, the literature has devoted little attention to understanding the reasons for the generalised failure of occupational regulations.[3]

In this paper, we study entry into the legal profession in Italy and we investigate potential mechanisms that can explain why occupational licensing so often fails at selecting the best professionals. We document that law school graduates with relatives who are already operating in the profession are more likely to pass the entry exam, regardless of their GPA in law school. In fact, the percentage of connected candidates passing the exam does not increase significantly with their GPA. For unconnected candidates, i.e. those without family ties among licensed lawyers, GPA in law school matters a lot for the probability of passing the entry exam, and only those with the highest grades display pass rates similar to connected candidates. In addition, we find that connected lawyers earn more than unconnected lawyers all else equal, especially at the lower end of the distribution of law school GPA. Our analysis shows that such earnings' advantage is substantially larger when young connected lawyers work in the same law firms as their family ties.

Due to the combination of these effects, positive selection on academic ability, as measured by GPA, is very limited. Among law school graduates in the lowest GPA decile, about 46% eventually become licensed lawyers compared to 50% in the highest decile. With the aid of a simple model, we simulate the selection process under the assumption that family connections were unimportant, both for the probability of passing the exam and for the earnings process. Our results show that occupational licensing would indeed produce strong positive selection on academic ability in this hypothetical scenario: the share of lawyers would decrease along the entire distribution of GPA, but the effect would be four times larger in the first than in the tenth decile.

One obvious explanation for these findings would be that connected lawyers learn the trade within the family and eventually accumulate higher occupation-specific human capital, even if they do not do well in law school. Although we cannot measure professional ability directly, we present several pieces of empirical

[2] Anderson et al. (2020) is the only paper that finds positive effects on quality of introducing regulation.

[3] Some even argue that in a world of online transactions, the prevalence of consumer ratings might make licensing redundant (Farronato et al., 2020).

evidence that appear to be at odds with this explanation. First of all, we document that, conditional on parental education, family connections have no predictive power on any available measure of human capital, most notably high school and law school grades. Moreover, we also show that connected lawyers are not more likely than their unconnected counterparts to work in large firms, to hold a position of responsibility in the workplace nor to move into the more competitive markets, which is what one would expect if they were more able professionals.

Based on these pieces of evidence, we believe that the differences between connected and unconnected professionals that we uncover in our analysis are more likely to arise because of favouritism and nepotistic practices rather than human capital accumulation within the family.[4]

Based on this interpretation, our analysis suggests that the interplay of poorly designed regulation and the intergenerational transmission of occupations can undermine the potential of licensing to generate the positive selection it is designed to create. This is a very important result for policy design, as occupational licensing affects about 20% of workers in the European Union and up to 30% in the United States (Koumenta and Pagliero, 2018; Kleiner and Krueger, 2013).

This paper offers a rationale for the failure of occupational regulation documented in so many countries and for so many professions. Although the mechanisms that we highlight clearly cannot apply to all possible contexts, we believe that they are common enough to be useful for policy design, at least in occupations that are highly persistent within families and subject to regulations that may favour nepotistic practices. This is certainly the case for lawyers, a very important profession that is regulated in Italy in much the same way as in most other countries: only graduates from 5-year law schools can enter the profession, conditional on an 18-month apprenticeship period and an entry exam which consists of both a written and an oral part. The long compulsory apprenticeship period, the partial anonymity of the exam, the presence of incumbent lawyers in the exam commissions, and the regulation of professional practices which make it difficult for young lawyers to attract new clients, are all factors that may naturally favour young entrants who already have connections with established professionals.[5]

As further discussed in Section 1.2.1, these institutional features are not unique to the Italian setting, nor to the legal profession. In particular, long apprenticeship periods, the involvement of incumbents in the entry process, and restrictions to

[4] Lentz and Laband (1989) also find no evidence of occupation-specific human capital accumulation within the family for medical doctors in the US.

[5] The professional code of conduct of Italian lawyers, like those of many other countries, fixes price floors and forbids commercial advertising, thus making it extremely difficult to attract clients. Some reforms of the system were attempted in the early 2000s but professional associations have been able to make them largely ineffective (Basso, 2009; Pellizzari and Pica, 2010; Orsini and Pellizzari, 2012).

prices and commercial practices are extremely common across many professions, especially liberal professions, and countries (UK Office of Fair Trade, 2001; Paterson et al., 2003; Pellizzari and Pica, 2010; Pellizzari et al., 2011). In addition, in countries with selective tertiary education systems such as the US, college admission is a crucial step in the selection process and it has been widely documented to favour connected candidates over unconnected ones, often regardless of quality (Broscheid and Teske, 2003; Cannings et al., 1996; Chetty et al., 2020).[6]

In Italy, the bar exam is administered by 26 local districts, one for each court of appeal. Despite being common to many other countries (e.g. Germany, Canada, US), this decentralisation presents some peculiarities that can be exploited for identification purposes. Specifically, each year the written exams of each district are marked by the commission of a different randomly selected district. Given the substantial variation in grading standards across the country, this setup allows us to identify the effect of connections on entry and earnings separately.

Our data combine university administrative records covering the universe of all law school graduates between 2007 and 2013 with the lists of all licensed lawyers, allowing us to know which graduates eventually become lawyers and when. In addition, all graduates are interviewed at the end of their university program as well as one, three, and five years after graduation. From these surveys, we obtain information on family background, apprenticeship, and earnings.

Crucial to our analysis is the measurement of family connections, which we implement following a now rather extended literature using surnames. We code graduates as connected if their surname appears at least once in the district register among lawyers who obtained their license at least 25 years before their (presumed) first bar exam attempt (Güell et al., 2018, 2015; Angelucci et al., 2010; Brollo et al., 2017; Buonanno and Vanin, 2017).[7] Of course, we acknowledge that this is an imperfect proxy and we include extensive robustness checks to investigate the implications of measurement error for our main results (see Section 1.7.1).

We are not the first to look into the relationship between occupational regulation and the quality or output of producers. Our data only allow us to measure input quality in terms of academic GPA. We acknowledge that this does not necessarily correspond to professional ability nor to service quality and we adapt the

[6]Even though most US jurisdictions do not require formal apprenticeship periods in order to sit the bar exam, the character and fitness requirements for admission are hardly anonymous. Candidates are required to disclose substantial personal, financial, and professional information. Other common law countries (as well as Israel) require aspiring lawyers to serve in articling positions under the supervision of senior members of the profession. Most European countries, such as France and Germany, also require extensive apprenticeship periods and vocational training after graduation from law school.

[7]For robustness, we also experiment with the total number of times one's surname appears in the local register *without accounting for differences in generation.*

interpretation of our findings accordingly. Nevertheless, we believe that our analysis still provides an important contribution. Contrary to standard theoretical predictions, most papers in this literature find no or even negative effects: our main contribution is to describe a potential mechanism that could explain this surprising and rather robust finding.

Carroll and Gaston (1981) in their exploratory analysis had already concluded that *"[..] there is [..] evidence from several professions and trades that indicates that restrictive licensing may lower received service quality. We know of no contrary findings[..]."*. More recently, Kleiner et al. (2016) could not find any detectable improvement in the quality of health services when licensing regulations for nurses became stricter. Similarly, Kleiner and Kudrle (2000) show that stricter licensing requirements for dentists result in higher prices with no significant improvement in quality.[8] Barrios (2018) exploits changes in the licensing requirements for accountants to find that *"[...] restrictive licensing laws reduced the supply [...] and increased rents to the profession without drastically improving quality [...]"*. Haas-Wilson (1986) and Kugler and Sauer (2005) show similar results for optometrists and physicians, respectively.

An important profession that has attracted a lot of attention is teachers, and once again there does not seem to be clear positive effects of regulation on quality. Angrist and Guryan (2008) investigate the introduction of state-mandated teacher testing in the US and find positive effects on wages but no effect on quality, measured by teacher qualifications. Larsen et al. (2020) examine the effect of stricter licensing requirements for teachers in the US and find an increase in the left tail of the quality distribution.

Anderson et al. (2020) is perhaps the only study to document a clear positive effect on quality. They examine the staggered introduction of licensing requirements for midwives across US states and find significant reductions in maternal and infant mortality.[9]

A recent significant advancement in this literature is Kleiner and Soltas (2022), who develop a sufficient statistics approach to assess the overall welfare cost or benefit of occupational licensing. They apply their methodology to a variety of occupations exploiting variations in regulations across US states and finding an overall welfare loss, suggesting that even if there were quality effects, they are more than offset by the welfare loss due to higher prices and lower supply.

This paper is also tightly connected to the literature on the intergenerational transmission of occupations. Lentz and Laband (1989) and Laband and Lentz

[8] Wanchek (2010) and Wing et al. (2005) also investigate occupational regulations for dentists but do not focus on quality.

[9] Deyo et al. (2020) also look at quality, but rather indirectly by studying the implications on crime and health of licensing massage therapists.

(1992) had already documented strong intergenerational persistence of professions for doctors and lawyers in the US, and rationalised this evidence with either nepotism or transmission of human capital within the family. Dunn and Holtz-Eakin (2000) and Bjorklund et al. (2012) further find similar results for general self-employment and capitalist dynasties, and Corak and Piraino (2011) even document that parents and children are often employed by the same employers. More directly related to our work, a recent literature documents sizeable intergenerational correlations of professional affiliations in Italy (Aina and Nicoletti, 2018; Mocetti, 2016; Mocetti et al., 2022; Raitano and Vona, 2021; Mocetti and Roma, 2021; Bamieh and Cintolesi, 2021). Compared to these papers, we link the intergenerational transmission of occupations to the effectiveness of licensing regulations by directly addressing the selection and quality of professionals.

Many other papers have looked at occupational licensing in a variety of professions, but without focusing specifically on quality. Most of these studies document an increase in costs for consumers and profits or rents for incumbents. This is the case for driving schools in France (Avrillier et al., 2010), lawyers (Pagliero, 2010, 2011), barbers (Thornton and Weintraub, 1979; Timmons and Thornton, 2010), and radiologists (Timmons and Thornton, 2008) in the US.

The rest of the paper is organized as follows. Section 1.2 describes the institutional setup of the legal profession in Italy. Section 1.3 presents our main data sources and how we combine them. The model that guides our empirical investigation is introduced in Section 1.4. The empirical implementation of the model and the results are discussed in Section 1.5. In Section 1.6 we present counterfactual simulations allowing us to quantify the role of various mechanisms in the process of selection into the legal profession. Section 1.7 contains a large battery of robustness checks. Section 1.8 concludes.

1.2 Institutional Background

The regulation of the legal profession in Italy is similar to many other countries. A government-issued license is required to offer legal services to clients and represent them in court. Only graduates from law schools, offered by either public or private universities, can obtain the license conditional on completing 18 months of compulsory practice and passing an entry exam.

The exam takes place in the courts of appeal, the second layer of the Italian judiciary system.[10] There are 26 such courts in the country, approximately one

[10] Beside the organisation of the lawyer entry examination, courts of appeal are mainly responsible for appeals against judgments issued by the ordinary courts, the first layer of the system.

per region, with the most populated regions having more than one. For simplicity, we will hereafter refer to these 26 courts of appeal as districts. The exam takes place once per year and consists of two parts. First, candidates sit a written exam that lasts three consecutive days. On the first day, they must write an opinion on a civil case, on the second day on a penal case, and on the third day, they write a judiciary act (e.g. court summon, complaint, succession, etc.). The dates and the texts of the exam are the same throughout the country, but each district has its own location and grading commission. Candidates sit the exam in the district where they did their apprenticeship.

The written tests are graded anonymously by the commission of a randomly assigned district. The randomisation is performed by the Ministry of Justice and is clustered within 5 groups of districts with similar sizes. It is also designed to avoid pairing, namely two commissions grading each other. As an example, in 2019 the group of largest districts, comprising Rome, Naples and Milan, had Naples grading Rome, Rome grading Milan, and Milan grading Naples. The outcome of the randomisation process is made public at or after the start of the written exams.[11]

Table 1.1 provides supportive evidence of the randomisation of marking districts. The Table reports results from a series of linear regressions with exogenous characteristics of law school graduates as dependent variables and a full set of dummies for grading districts as explanatory variables, also conditional on district of origin and group×year fixed effects.[12] Proper randomisation should imply that the coefficients on the marking district fixed effects are all zero. For each dependent variable, the Table reports the F-test – and corresponding p-value – for the joint significance of all the marking district fixed effects (columns 1 and 2), the share of statistically significant fixed effects at the levels of 90%, 95%, and 99% (columns 3,4 and 5). While the F-test almost never rejects the null that all fixed effects are jointly equal to zero, the number of significant marking district dummies is very small and consistent with the randomisation taking place within small groups.

For a variety of reasons ranging from differences in local norms to idiosyncrasies in composition, commissions in different years and districts may be more or less lenient. Hence, the random assignment of grading commissions generates exogenous variation in the probability of passing the entry exam and it is a crucial element of our empirical strategy. Specifically, we assume that the randomly as-

[11] For example, in 2019 the written exam took place on December 10-11-12 and the grading commissions were announced on December 10. In 2018 the written exam took place on December 11-12-13 and the grading commissions were announced on December 21.

[12] The total number of marking district fixed effects is 26. The data are described in Section 1.3.

Table 1.1: Balance Table

Dep. Variable	Joint F-test	p-value	Share of Stat. Sig. FEs[a]		
			90%	95%	99%
	(1)	(2)	(3)	(4)	(5)
1=connections[b]	1.638	0.023	0.115	0.038	0.000
GPA[c]	1.588	0.031	0.038	0.038	0.000
High school grade[d]	1.906	0.004	0.115	0.077	0.000
1=graduate parent[e]	2.370	0.000	0.115	0.077	0.038
1=parent(s) in high-ranked occup.[f]	5.065	0.000	0.077	0.038	0.000
1=female	1.070	0.369	0.192	0.077	0.038
Age at graduation	1.523	0.046	0.038	0.000	0.000

[a] There are 26 grading district fixed effects. Each regression includes also fixed effects for district of origin and group×year. [b] At least one person (25y+ older) with the same surname appears in the local register at the (expected) time of sitting the bar exam. [c] Grade point average for all graded exams taken over the five-year law school program, weighted by academic credits and standardised within university. [d] Standardised over the sample. [e] At least one parent with a university degree. [f] At least one parent employed as a professional, entrepreneur, or executive manager.

signed marking commission affects the probability of entering the legal profession, but is excluded from the earnings process, thus allowing the separate identification of the role of connections on entry and earnings.

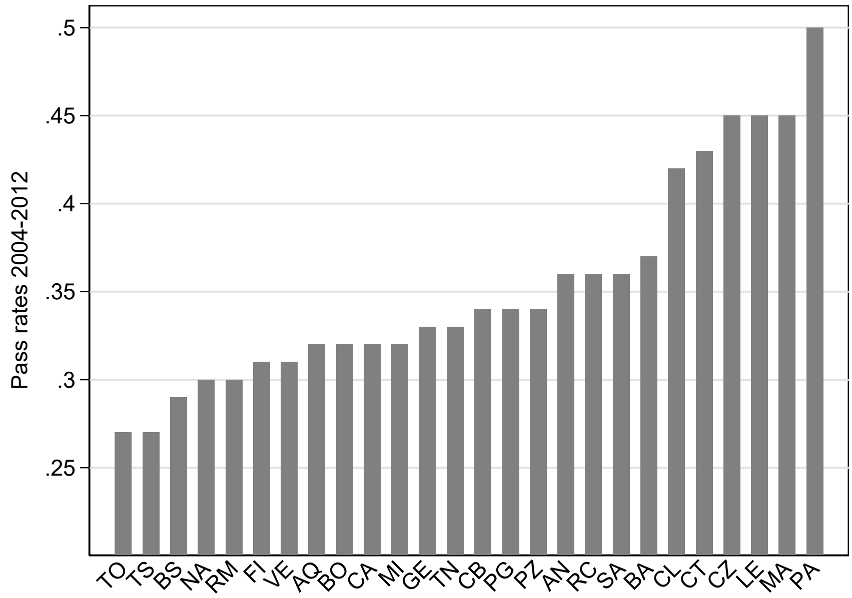
Figure 1.1 shows the average unconditional pass rates at the entry exam – both written and oral (details on the oral part below) – over the period 2004-2012 in every district. On average 34% of candidates eventually pass, but there are very large differences across the country, ranging from 27% in Turin (TO) to 50% in Palermo (PA).[13]

Every year several thousand candidates attempt the written exam: on average, over 30,000 candidates participated every year during the period of our analysis. Hence, the grading process is long, usually lasting around 6 months. The written exam takes place at the same time for all districts, normally at the beginning of December. The results are published during the summer and successful candidates are then admitted to the oral exam. The interviews happen in alphabetical order, starting with a letter that is randomly drawn by each district. The oral examinations are scheduled independently in each district and start as soon as the results of the written exams are available. The interview usually takes about one hour and candidates are immediately notified about the outcome. In most districts the interview calendar spans from September to December and, eventually, the entire process is completed just before a new round of written exams begins.[14]

[13]This descriptive finding may simply reflect heterogeneity in the pool of applicants across districts. We will later show conditional evidence of cross-district heterogeneity in grading standards.

[14]It is common for candidates who successfully passed the written exam but are waiting to

Figure 1.1: Pass rates at the bar exam



Source: Buonanno and Pagliero (2018).

Given the 18 months of practice and the length of the examination process, young lawyers obtain their license approximately 2.5-3 years after graduation, unless they fail the exam (either the written or the oral part), in which case the process may take substantially longer. The entry exam can be retaken any number of times.[15] Figure 1.2 summarizes the entry process into the profession, from the moment of graduation to the final occupational outcome. Candidates who fail the exam can either retake or, in many cases, choose to enter a different occupation, often as legal consultants in private firms.

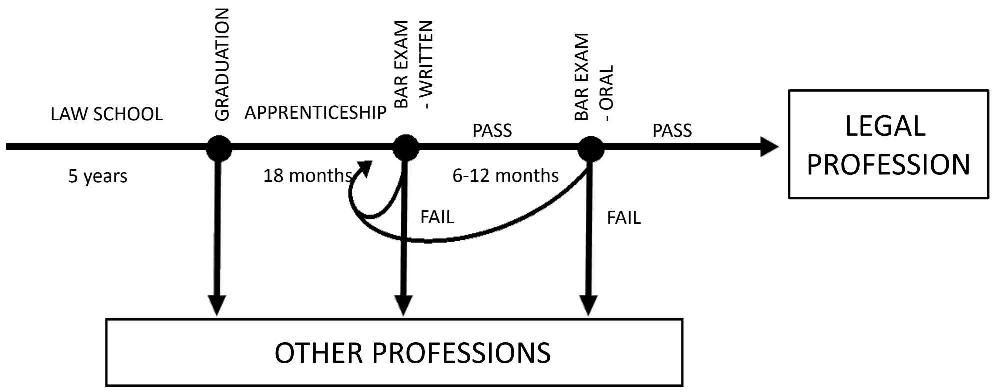
Candidates who successfully pass both the written and oral exams can then register with the local bar associations and operate in the corresponding local market. There exists one bar for each ordinary court, the lower level of the judiciary system, corresponding approximately to administrative provinces. Hence, there are multiple bars (on average about 5) in each district. In total, there currently are 139 local bar associations that are responsible for enforcing the professional code of conduct and organising training for their associates.[16] Lawyers are only allowed to represent clients in the ordinary and appeal courts outside their local bar if they pair with a local lawyer, but they can freely choose to transfer to any

take the oral part to enroll in the written exam of the following year, to be ready to sit it again in case they fail the interview.

[15]The completion of the 18-month training period is valid only for 5 years.

[16]The number of local bars has varied slightly over time due to the separation of a few large ones and the re-aggregation of smaller ones.

Figure 1.2: Timeline of the licensing process for Italian lawyers



bar in the country at any time of their career. Registered lawyers can only work as self-employed professionals and cannot be dependent employees in the private sector (exceptions are possible in the public sector).

The local bar associations play an important role also in the organisation of the entry exam as they nominate three of the five components of the local exam commission. The other two members are a judge (usually retired) and a university professor and they are nominated by the Ministry of Justice. The president of the commission must be chosen among the three lawyers. The local commissions are responsible for the logistics of both the written and the oral examinations, they mark the written exams of the candidates of the randomly assigned district and they carry out the oral interviews of the local candidates who passed the written exam.

A central commission with the same composition is created by the Ministry of Justice and it is responsible for preparing the written exam questions, defining grading criteria, and overseeing the entire examination process. All commissions change every year.

1.2.1 International comparison

The overall structure of the licensing process for lawyers is quite similar across most industrialised countries, Italy included. Virtually everywhere aspiring lawyers need to graduate from law school, complete some compulsory vocational training and go through an exam-based admission process.[17]

The characteristics of the law degrees which give access to the vocational training are usually also highly uniform across countries. Within the European Union, agreements are in place to allow the automatic mutual recognition of degrees, and systems of minimum requirements determine the validity of degrees across a broader set of countries. In most Western countries, access to the legal profession requires the equivalent of 4 to 5 years of tertiary education. The subsequent professional apprenticeship is usually organized in collaboration between universities and the state with slight differences across countries. In most common law countries (e.g., UK, Canada, Australia), the young graduates go through a compulsory articling period, during which they train directly with senior members of the bar.[18]

In the US, graduates must enroll in a post-graduate American Bar Association (ABA) accredited law school, which includes some vocational training. Alternatively, some states accept work periods within the court system as an alternative to law school. Overall, even though a mandatory articling period is seldom required in the US, the system encourages aspiring lawyers to obtain on-field training through pro-bono programs, clerical work and supervised "Public Service Requirements" (now compulsory in certain law schools). In France, law graduates must obtain a state-administered vocational degree (*certificat d'aptitude à la profession d'avocat (CAPA)*), which normally requires attending a post-graduate law school (including both academic and vocational training), with entrance through a competitive examination. Germany requires two state-administered exams to enter the legal profession: a first one after university which allows successful candidates to qualify for two years of compulsory training period (*Referendariat*), and a second one after successful completion of the training. Israel follows a similar system, with two state-administered examinations (one after university and one after the vocational training), and one year of articling, which is accessible conditional on passing the first exam successfully. Other countries requiring a compulsory articling period after obtaining a law degree at university are Singapore (six months), Spain (two years), Poland (the duration depends on the specialization), Iran (18 months), Finland (four years), Denmark (three years), Japan (one year), India (two years).[19] All these countries then require passing a state-administered exam,

[17] One notable exception is the state of Wisconsin in the US, where individuals who obtained a degree in Law from an American Bar Association accredited school in the state may be admitted to the state bar through diploma privilege.

[18] Most common law countries require different vocational training and state-administered exams depending on whether a candidate wishes to pursue a career as a barrister or as a solicitor. Even though this difference does not exist in Italy, it does not change the admission mechanism substantially, as both careers require university training, articling, and passing a state-administered exam.

[19] Japan's vocational training takes place after sitting a nationally administered exam, which has the lowest success rate in the world, around 22%.

which upon successful completion allows an individual to practice law.[20]

The written exam is always anonymous, however many countries require some sort of oral examination (Italy, Germany), trial examination (Finland, Australia), or character and fitness requirements (US), which, by their nature, cannot be anonymous. Furthermore, most countries organize their local bar associations in a way similar to Italy, which are responsible for administering the final admission exam.

Given the information we collected, we observe that most countries broadly follow a three-step procedure to regulate access to the legal profession: a tertiary degree, a compulsory apprenticeship, and a state-administered exam, which is seldom completely anonymous. The similarity of the regulations across countries suggests that the mechanisms that we highlight in this study, notably the role of the inter-generational transmission of occupation, are likely to be present in many other contexts.

1.3 Data and descriptive evidence

We combine various data sources to follow several cohorts of Italian law school graduates over the first 5 years of their careers. The starting point is an administrative dataset covering (almost) the entire universe of university graduates in Italy.[21] The dataset is constructed and maintained by AlmaLaurea, a consortium of Italian universities sharing their administrative records for research purposes and offering placement services to both graduates and employers.

In addition to maintaining the administrative data, the consortium also runs a series of regular surveys of all graduates. A first survey takes place right before they graduate and collects information about the students' backgrounds, opinions about the university experience, and expectations about their professional careers. Then, students are interviewed again one, three, and five years after graduation to collect information about their labour market status. Almost all students fill in the survey at graduation, which is required to obtain their diplomas. The response rates of the other surveys are very high: on average around 80% at one year, 75% at three years, and 70% five years after graduation.

For this study, we focus on students who graduated from law school between 2007 and 2013. Before 2007, the follow-up surveys were only administered to those

[20] The level at which an individual can practice may vary from country to country: in Italy, separate training and competitive examinations must be undertaken to practice as a judge or a notary. Common law countries usually distinguish between barristers and solicitors, and countries such as Poland and Hungary allow for different specializations.

[21] All 50 public universities offering law degrees are included. Five of the nine private universities offering law degrees are included.

who graduated in the summer session (that is, about one-third of graduates) and data about only 49 out of the 76 participating universities was available, whereas the 5-year post-graduation surveys for students who graduated after 2013 have not yet been released.

Administrative records include high school type and marks, the university the students graduated from, GPA, graduation grades, age and gender. Important survey information comprises employment status and wages, parental education and parental occupation, scholarships, experiences abroad, proficiency in foreign languages and computer skills.

To identify graduates who eventually enter the legal profession, we match the main dataset with the official registers of licensed lawyers in the entire country. Local bar associations are responsible for publishing and maintaining the lists of licensed professionals in their jurisdictions and most of the associations make them available on their websites. We have collected all the registers from November 2017 to January 2018 and we observe the names, surnames and unique tax identification numbers of all the associates. We use this information to match our main dataset of law school graduates with the lists of licensed lawyers, allowing us to identify those who eventually entered the legal profession, when and where.[22]

This information might suffer from some inaccuracy. For example, some individuals might register and then unregister shortly after if they choose to leave the profession.[23] Given the high cost of entering the profession, we expect this to happen rarely. The opposite source of error is also possible, namely individuals who become lawyers more than 5 years after graduation and are not recorded as licensed professionals in our data. We also expect this to happen rarely, essentially only for candidates who fail the exam several times or who try a different career path first.

We further complement our data with measures of connection with the profession based on surnames. For all graduates, both those who eventually work as lawyers and those who do not, we compute how frequently their surnames appear in the local register (or in others). We assign the local district corresponding to the location of the university from which they graduate for both lawyers and non-lawyers.[24]

Of course, this is an imperfect measure of connectedness. There can be fam-

[22]The matching has been performed for us by AlmaLaurea and we only have access to the matched anonymised version of the final dataset, which we can only access on the consortium's premises.

[23]It is also unclear whether it would be correct to classify them as lawyers. Ideally, we would like to consider them as successful candidates when we look at the likelihood of passing the exam, but change their status to non-lawyers when we look at their earnings.

[24]We have also experimented with an alternative definition which assigns to lawyers the district where we observe them registered. The results are largely unaffected.

ily ties not sharing the same surname, like one's mother and her relatives, and, conversely, individuals sharing the same surname may not be connected to one another. We know from the literature that, given the usual Western conventions for surname transmission (and Italy is no exception), the second source of error is likely to be very small because the vast majority of individuals hold surnames that are highly infrequent. Hence, the probability that any two individuals with the same surname are linked by some family tie is extremely high.[25] It remains possible, however, that we fail to capture some connections because they do not share the same surname. In Section 1.7.1, we simulate various scenarios of mismeasurement to show that, under most assumptions, the degree of error must be extremely large to overturn our main findings, at least qualitatively.

In addition, we also observe the professional coordinates of each lawyer from public registers – postal addresses, emails and phone/fax numbers – allowing us to identify those who are likely working in the same law firm. More specifically, we assume that any two lawyers reporting the same phone number or fax address or postal address work in the same firm.[26] We use this information to compute the size of the law firms that appear in our data (recall that from the professional registers we observe all active professionals) and also to identify young connected lawyers who work in the same firms as their connections.

Finally, we collect information on the distribution of surnames in each district from tax records. Specifically, we compute the number of times each surname appears in the tax records of each district and we use this information to control for the incidence of each surname in the underlying population.[27]

In our empirical analysis, we estimate several equations and, due to missing values and survey non-response, the number of observations available for each of them varies. For comparability purposes, in our main analysis, we restrict the sample only to the observations that can be used for all equations, but in Section 1.7.3 we replicate all our estimates to show that the results are largely unaffected by this sample selection.

[25](Güell et al., 2018) compute that in Italy the probability of two people taken at random being family members, conditional on having the same surname, is 0.1838, which is about 2000 times higher than the unconditional probability.

[26]Multiple firms may have the same postal address if they are located in the same building. We manually checked (by searching their websites) the largest resulting studios to confirm that indeed the coordinates refer to a single firm. Many younger lawyers do not provide fax numbers and 9.97% of those who provide a phone number give a mobile line. In the end, we are able to retrieve addresses for 240,727 out of 240,957 registered lawyers (99.9%), telephone numbers for 220,438 lawyers (91.5%), and fax numbers for 192,609 lawyers (79.9%). We use the more precise information on landline phones, fax and email to validate matching based on postal addresses, and we conclude that misclassification is a minor problem.

[27]We extract this information from the same data used in Güell et al. (2018). The population appearing in the tax records is not exactly identical to the total population, but Güell et al. (2018) show that it is a reasonable approximation, especially regarding the adult population.

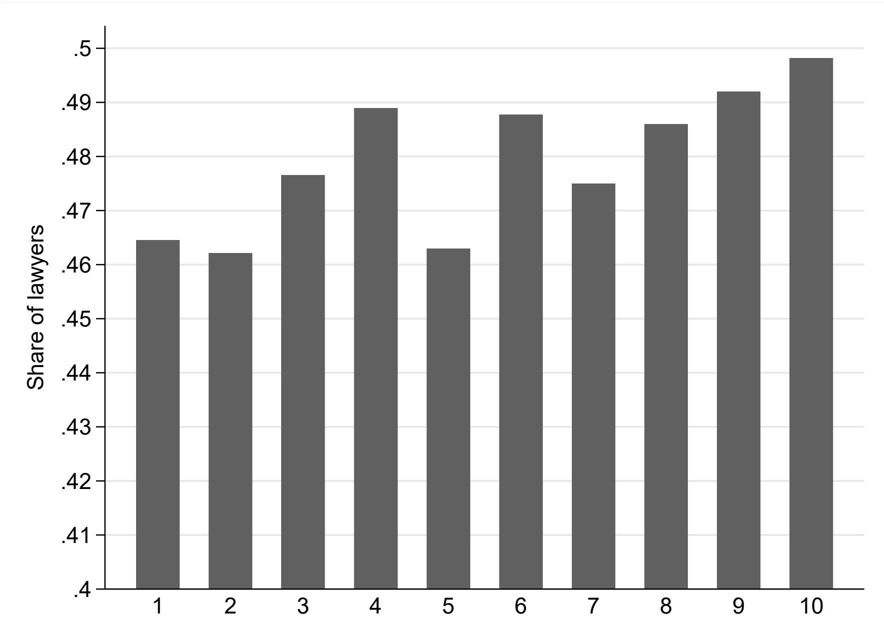
Table 1.2: Descriptive statistics

	Full sample	Lawyers[a]	Non-lawyers[a]
1=female	0.63	0.62	0.64
	(0.482)	(0.485)	(0.479)
High school grade[b]	-0.00	0.03	-0.02
	(1.000)	(0.984)	(1.014)
GPA[c]	0.00	0.02	-0.02
	(0.999)	(0.977)	(1.003)
1=connected[d]	0.58	0.62	0.53
	(0.494)	(0.485)	(0.499)
Number of connections[e]	4.14	4.54	3.76
	(12.26)	(13.16)	(11.36)
1=graduate parent(s)[f]	0.38	0.40	0.37
	(0.486)	(0.489)	(0.483)
1=parent(s) in high-ranked occup.[g]	0.44	0.46	0.41
	(0.496)	(0.498)	(0.492)
1=apprenticeship[h]	0.88	1.00	0.77
	(0.323)	(0)	(0.419)
Log earnings[i]	5.82	5.96	5.69
	(2.600)	(2.266)	(2.868)
Observations	24260	11629	12631

[a] Graduates who appear or not in some local register of lawyers in 2017/2018. [b] Standardised over the sample. [c] Grade point average for all graded exams taken over the five-year law school program, weighted by academic credits and standardised within university. [d] At least one person (25y+ older) with the same surname appears in the local register at the (expected) time of sitting the bar exam. [e] Number of persons (25y+ older) with the same surname appearing in the local register at the (expected) time of sitting the bar exam. [f] At least one parent with a university degree. [g] At least one parent employed as a professional, entrepreneur, or executive manager. [h] Graduates who self-reported having started a legal apprenticeship in at least one post-graduation survey (one, three and five years after graduation) or who are registered as apprentices in the official lawyer registry. [i] Self-reported earnings five years after graduation (in Euros 2015).

Table 1.2 reports some basic descriptive statistics for the 24,260 individuals in this common sample, broken down by those who eventually enter the legal profession and those who do not. Law school attracts over 60% female students and a little majority of them eventually end up not practising as licensed professionals. The descriptive statistics suggest some minor positive selection on academic ability into the legal profession, both looking at high school and university grades.[28] Over half of the graduates have some connection with the profession and the incidence of connections is substantially higher among those who eventually enter the profession. These students are also slightly more likely to come from educated and affluent families, which we measure with parental education and occupation. The data suggests that most law school graduates attempt to enter the legal profession: 88% of them start the apprenticeship and eventually 77% of those who do not become licensed lawyers report having started an apprenticeship. Earnings five years after graduation are already significantly higher for lawyers than non-lawyers by about 10% of a standard deviation.

Figure 1.3: Share of lawyers by deciles of GPA



We conclude this section by presenting some descriptive evidence on the selection of lawyers into the profession and the role of family connections. Figure 1.3 plots the share of law school graduates in our sample by decile of the distribution of GPA. This plot shows that some positive selection on academic ability

[28]To account for differences in grading standards, we have standardised GPA to have mean zero and standard deviation equal to one within each university. High school final grades are instead standardised across the entire sample because they are attributed via a common national exam.

does take place, but it is quite limited.[29] About 46% of graduates in the first and second deciles access the profession and this share barely reaches 50% at the very top of the distribution. Section 1.4 will go beyond this descriptive evidence by accounting for the process through which individuals self-select into the profession, allowing us to quantify the importance of family ties in limiting the amount of positive selection.

Figure 1.4: Share of connected individuals by occupation and GPA deciles

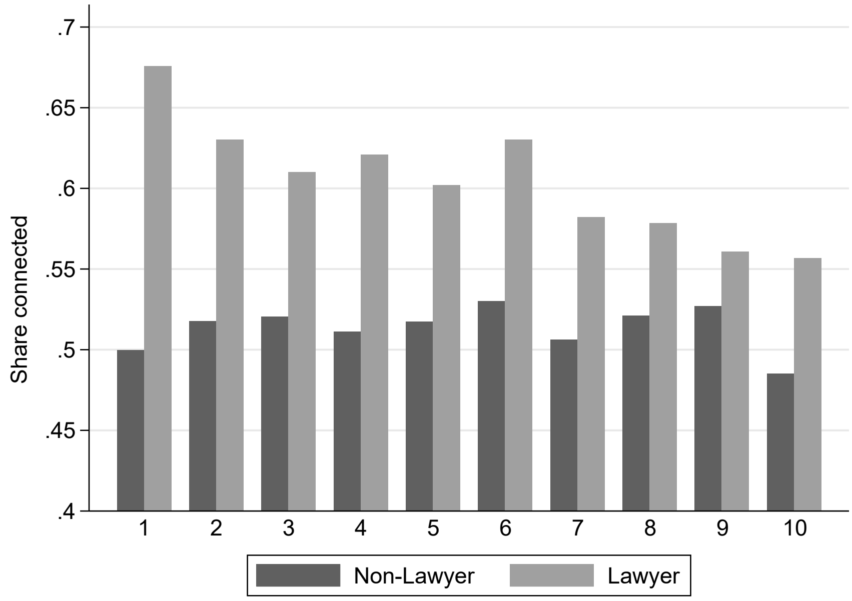


Figure 1.4 shows instead the share of connected individuals in our sample, broken down both by deciles of GPA and by the groups of those who eventually become licensed lawyers and those who do not. Several notable facts emerge from this figure. Family connections are much more frequent among lawyers than non-lawyers throughout the distribution of GPA, but differences are much larger at the bottom. Among the least able graduates, almost 70% of those who eventually enter the legal profession have some family member who is already a licensed lawyer. For those who end up in a different occupation, this figure is about 52%. At the top of the distribution of GPA, the difference is around 6 percentage points (56% versus 50%). The share of connected lawyers evidently declines with GPA whereas it appears to be rather flat for non-lawyers.

Taken together, Figure 1.3 and Figure 1.4 suggest that family connections

[29] Becoming a judge or a notary is another potential occupational outcome for law school graduates. The selection processes are much longer than for lawyers, hence graduates choosing these careers appear as non-lawyers in our data. Nevertheless, the number of available seats for both judges and notaries is low and fixed by law. It is unlikely that considering these professions separately might change our findings.

might interfere with the selection process and explain, at least in part, the mild positive selection on GPA that we detect in the raw data.

1.4 A model of occupation choice with regulated professions

In this section, we present a statistical model of selection into the legal profession. The model is developed with the explicit purpose of being implemented empirically. Hence, we tailor it to maintain proximity with our data, but the model remains quite general given the similarity of the Italian legal profession with many other regulated occupations around the world (see Section 1.2.1).

The model consists of a sequential process, mimicking the scheme in Figure 1.2. First, individuals accumulate human capital and we allow this process to take place both in school, college and at home. Second, at the end of college, agents make occupational choices, namely whether they want to try entering the regulated legal profession or not. Those who choose the legal profession need to do an apprenticeship and pass the entry exam, whereas the others can immediately start producing earnings in another non-regulated occupation. Agents who choose the legal apprenticeship but fail the entry exam work in some other non-regulated occupation, whereas those who successfully pass become lawyers and generate earnings from professional practice. The following paragraphs describe how we model each of the steps, starting from earnings and moving backwards in the sequence of events.

1.4.1 Earnings

Once the entire process of occupational selection has played out, a generic agent i can be employed as a lawyer or as a non-lawyer and her earnings are determined as follows:

$$Y_i = \begin{cases} y^L(A_i, S_i, N_i, G_i, X_i) + u_i^L & \text{if working as a lawyer} \\ y^0(A_i, S_i, N_i, G_i, X_i) + u_i^0 & \text{if working as a non-lawyer} \end{cases} \tag{1.1}$$

where A_i is general ability, S_i is occupation-specific (legal) ability, N_i is a measure of connection with the legal profession and G_i is parental human capital. With all the necessary caveats and caution, we proxy A_i with high school grades, S_i with GPA in law school and N_i with the presence of older licensed lawyers with i's same surname in the same district. We allow the human capital of the parents to have a direct effect on earnings beyond ability and connections to account for dimensions of ability that are not captured by A_i and S_i. We measure G_i with parental education. X_i is a set of additional controls including gender, age at graduation and dummies for graduation years, district and university. In order to make our measure of connectedness comparable across individuals with more or less popular surnames, we also condition on the log number of individuals with own surname in the district and log population size of the district.

1.4.2 Compulsory apprenticeship and bar exam

After graduation, individuals make occupational choices. Those who choose the legal profession must first complete the compulsory apprenticeship and then pass the bar exam. Individuals who, instead, choose other professions can start producing earnings right after graduation.

Let us start by describing how we model the probability of passing the bar exam. We assume that the overall performance at the exam is a function $p(\cdot)$ of general and occupation-specific skills, individual and family characteristics, and a random shock ϵ_i, that is realised only on the day of the exam (e.g. luck, fatigue, anxiety, etc.):

$$p(A_i, S_i, G_i, X_i) + \epsilon_i \tag{1.2}$$

The agent passes the exam if her performance is above a given threshold, which we allow to vary according to the strictness of the grading district and on one's connections. Recall that the exam consists of both a written and an oral part. The written part is marked by a randomly selected district and in Section 1.2 we have documented the large heterogeneity in grading standards across districts (see Figure 1.1). Hence, being randomly assigned to a lenient or strict district may substantially affect the probability of passing the exam. Next, the oral part takes place in one's local district and it obviously cannot be anonymous. Hence, nepotistic practices may emerge at this stage of the process and connected candidates may be more likely to pass.

We define the minimum performance threshold to pass the exam $t(R_r, N_i)$, where R_r is the grading standard of district r and N_i is our indicator of connectedness for agent i. District r is the district grading i's written exam, which varies both by district and over time. We do not have exact information on the year when the individuals in our sample took the exam and some of them might have done it multiple times. However, even if we had this information it would be quite difficult to interpret it because both the decision to postpone the exam and failing it are clearly endogenous to the processes we are modelling. Hence, we simply define the grading district r as the district that was randomly assigned to grade the written exams of i's own district in the third year after i's graduation. Considering that the exam takes place only once per year in December, that most graduations happen in Spring/Summer and that the apprenticeship lasts a minimum of 18 months, for the vast majority of individuals the third year after graduation is the first time when they could theoretically take the exam.

Eventually, the probability of passing the bar exam is defined by the following event:

$$p(A_i, S_i, G_i, X_i) + \epsilon_i \geq t(R_r, N_i) \tag{1.3}$$

At graduation, agents make their occupational choices taking into account the probability of passing the exam, the cost of the apprenticeship period and expected future earnings, in the legal profession or in other occupations. For simplicity, we assume that apprentices are not remunerated and we define the cost of the apprenticeship as a function of the family's socioeconomic status.[30] The intuition is that affluent parents are better able to support their children during this relatively long period with no or little income. In our data, we do not observe family income and we proxy socioeconomic status with parental occupation, namely whether one or both of the parents work in high-paying occupations, such as professionals, managers and entrepreneurs.[31] Let this indicator be W_i and the cost of the apprenticeship $C(W_i)$.

We further assume that the idiosyncratic component of earnings u_i in equation (1.1) is realized only upon entering the labour market and that its conditional mean is zero, i.e. $E(u^J|A_i, S_i, N_i, G_i, X_i) = 0$ with $J = \{L, 0\}$. Then, agent i chooses to start an apprenticeship and eventually sit the exam if:

$$P\left[\epsilon_i \geq t(R_r, N_i) - p(A_i, S_i, X_i)\right]\left[y^L(A_i, S_i, N_i, G_i, X_i) - y^0(A_i, S_i, N_i, G_i, X_i)\right] + v_i \geq C(W_i) \tag{1.4}$$

where v_i is an idiosyncratic preference component that is unobservable to the econometrician but known to the agent.

Eventually, the probability of becoming a lawyer can be computed as the product between the probability of starting an apprenticeship (equation (1.4)) and the probability of passing the bar exam, conditional on having started an apprenticeship (equation (1.3)).

[30] There is no legal requirement to remunerate apprentices and, in practice, few of them receive a salary.

[31] This follows a relatively standard definition of social groups that is also adopted by the Italian National Statistical Institute (ISTAT) (ISTAT, 2017).

1.4.3 Academic ability

One important innovation of our data is the availability of a measure of academic ability in legal matters, namely GPA in law school. We model the formation of such ability as follows:

$$S_i = s(A_i, N_i, G_i, W_i, X_i) + e_i \tag{1.5}$$

where all variables have the usual meaning and e_i is an idiosyncratic error term. It seems natural to allow generic ability, A_i, to influence the formation of S_i. Importantly, we also allow connectedness to affect academic ability, as one can learn from parents or other relatives with experience in the profession.

Our main interest in the estimation of equation (1.5) is related to the role of family connections (N_i). We want to investigate whether and to what extent law school students with relatives who are already active in the legal profession perform better than their unconnected peers.

Given the nature of our data, which only contains measures of ability at the end of law school, we abstract from the additional learning that might take place during the apprenticeship. In Section 1.5.3 and Appendix A, we discuss the implications for our analysis of a more complex process of skill formation that takes place also during the apprenticeship period.

Identification of equation (1.5) might be complicated by omitted variables, most notably innate ability. We believe that the problem is relatively minor in our setting because the explanatory variables that we include in these equations are unlikely to be endogenous.[32] At a minimum, equation (1.5) can be identified under the assumption that, once controlling for general skills via high school grades, innate ability would have no direct effect on specific skills, which is a commonly used assumption for proxy variables. We maintain this assumption also for the identification of all the other equations, but we return to its implications in Appendix A.

1.5 Empirical implementation and results

Conditional on imposing functional form and distributional assumptions, our data allows us to estimate all the equations of the simple model presented in the previous section. The resulting estimates are interesting in their own right but

[32] Perhaps the one variable that might be the most problematic is W_i, the indicator for having at least one parent working in high paying occupations. However, given that we proxy the socio-economic status of the family with predetermined parental occupation, we find it unlikely that this indicator could be affected by the children's innate ability.

taken together and interpreted through the lenses of the model, they also allow us to run simulations where we change a number of structural features and analyse their implications for the selection of professionals.

In this Section, we present our main estimates and we leave the simulations to the next section (Section 1.6). The main results are produced using the restricted sample of observations that are available for all equations (see Table 1.2), thus avoiding issues with the composition of the sample when comparing results across equations. In the robustness checks of Section 1.7.3, we show that these main findings are robust to changing samples across equations.

We estimate most equations of our model separately. Of course, it is possible to also estimate them jointly and we do it when necessary for identification purposes. For example, we jointly estimate the earnings equation (1.1) in a switching regression model, where we use dummies for the randomly assigned grading district as exclusion restrictions (parental occupation is an additional exclusion restriction). Otherwise, we prefer to limit the number of required distributional assumptions and estimate equations separately.

We conclude this section (Section 1.5.4) with a set of results that help guide the interpretation of the overall findings, especially regarding the role of occupation-specific human capital and its intergenerational transmission.

1.5.1 Academic ability

We start with equation (1.5), which describes how legal academic ability is formed. We assume linearity and we estimate it by simple OLS:

$$S_i = \beta_0 + \beta_1 A_i + \beta_2 N_i + \beta_3 G_i + \beta_5 W_i + \beta_4 X_i + e_i^S \tag{1.6}$$

Results are reported in Table 1.3. Perhaps not surprisingly, law school GPA is positively associated with both high school graduation marks and parental education. More important for the purpose of our paper is the lack of a meaningful and statistically significant association between law school GPA and our indicator of connectedness with the legal profession. If anything, the results in Table 1.3 indicate that law school graduates with at least one relative in the local register have slightly lower GPA, although the estimated coefficients do not reach conventional levels of statistical significance.

In our simplest specification (column 1), the estimated β_2 is equal to -0.014 (1.4% of a standard deviation, given the standardisation of GPA) with a standard error of 0.014, implying that a standard one-sided test assigns a probability of 83.9% to the coefficient taking any non-positive value.

The following columns of Table 1.3 investigate whether this finding might be

Table 1.3: Occupation-specific human capital

Dep. variable= GPA^a	(1)	(2)	(3)	(4)
High school grade^b	0.402***	0.402***	0.402***	0.402***
	(0.006)	(0.006)	(0.006)	(0.006)
1=connections^c	-0.014	-	-	-
	(0.014)			
1= few connections^d	-	-0.012	-	-
		(0.014)		
1= many connections^d	-	-0.028	-	-
		(0.020)		
Number of connections	-	-	-0.001*	-0.001
			(0.001)	(0.001)
Number of connections^2	-	-	-	-0.000
				(0.000)
1=female	0.098***	0.098***	0.098***	0.098***
	(0.011)	(0.011)	(0.011)	(0.011)
1=graduate parent^e	0.126***	0.126***	0.125***	0.125***
	(0.011)	(0.011)	(0.011)	(0.011)
Observations	24,260	24,260	24,260	24,260

^{a} Standardised within university. ^{b} Standardised over the sample ^{c} 1=some connections; 0=no connections. ^{d} few = 1-3 ; many = 4+. ^{e} At least one parent with university degree.
All regressions include fixed effects for university, district, year of graduation, log size of district and log name frequency in district. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

due to nonlinearities in the relationship between GPA and connectedness. Column 2 categorises connections into three broad groups: no connections (the baseline), few connections (1 to 3), and many connections (4 or more). Columns 3 and 4 further look at the linear number of connections and its square. None of these specifications points towards a positive effect of connectedness on GPA.

One may argue that having a relative in the profession may help develop a set of skills that are not necessarily captured by performance in university exams. Unfortunately, we do not have direct measures of the most obvious suspects, such as the ability to speak in public or to inspire confidence. However, the surveys include a variety of variables that should capture other dimensions of ability, such as certified knowledge of foreign languages and computer skills, whether the person engages in volunteering activities, or whether she has done a study exchange abroad. In Table 1.4 we report the estimates of regression equations like equation (1.6) but with each of these indicators as dependent variables.

While students with higher high school grades are more likely to hold certifications of proficiency in both foreign languages and computer skills, and are

Table 1.4: Additional measures of human capital

	(1)	(2)	(3)	(4)
Dep. Variable	Languages[a]	Computer Skills[b]	Volunteering[c]	Study Exchange[d]
High school grade[e]	0.069***	0.019***	0.007**	0.001
	(0.003)	(0.003)	(0.003)	(0.002)
1=connections[f]	0.016*	-0.007	-0.003	0.004
	(0.008)	(0.007)	(0.009)	(0.006)
1=female	0.013**	-0.011**	0.004	-0.015***
	(0.006)	(0.005)	(0.006)	(0.005)
1=graduate parents[g]	0.093***	-0.015***	0.035***	0.080***
	(0.006)	(0.005)	(0.006)	(0.005)
Observations	21,983	22,655	21,448	24,260
Mean of dep. variable	0.312	0.195	0.257	0.131

[a] Dummy equal to 1 if the respondent holds an internationally-recognized language certificate (e.g. TOEFL). [b] Dummy equal to 1 if the respondent holds the "European Computer Driving License" (ECDL). [c] Dummy equal to 1 if the respondent participates in volunteering activities. [d] Dummy equal to 1 if the respondent has spent a study period abroad (e.g. Erasmus). [e] Standardised over the sample. [f] 1=some connections; 0=no connections. [g] At least one parent with university degree.
All regressions include fixed effects for university, district, year of graduation, log size of district and log name frequency in district. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

also more likely to volunteer, the coefficient on connectedness is positive and (marginally) significant only for foreign languages. The magnitude of the effect is, however, very small as it amounts to a 5.1% change in the baseline probability to hold a certification of proficiency in a foreign language. The point estimates of the coefficients on connectedness for computer skills, volunteering, and study exchanges are all non-significant and very small.

One interpretation of the findings in this section is that, as far as Italian lawyers are concerned, the accumulation of occupation-specific human capital within the family is limited or even absent. Lentz and Laband (1989) find similar evidence for medical doctors in the US. Of course, it remains possible that connected individuals accumulate more occupation-specific human capital at later stages, for instance, because they have access to better training opportunities during the apprenticeship period. We address this concern in Section 1.5.4, where we present various pieces of evidence that help guide the interpretation of our results. For example, we show that connected and unconnected individuals do their apprenticeships at law firms of comparable quality. We also show that conditional on GPA, connected and unconnected lawyers are equally likely to work in large firms, hold positions of responsibility in their firms, or move to the largest and most remunerative markets. If connected lawyers had higher occupation-specific human capital for a given GPA, one would expect them to detect differences in at least some of these outcomes. Hence, we believe that our findings are more consistent with the interpretation that the intergenerational transmission of legal ability is limited.

1.5.2 Compulsory apprenticeship and bar exam

Our data allows us to identify both those graduates who at some point during their first 5 years after graduation started an apprenticeship and those who eventually pass the bar exam and register as lawyers. Hence, we can estimate both equation (1.3), which describes the probability of passing the exam (both written and oral), and equation (1.4), which describes the probability of starting the apprenticeship. In both cases, we need to make distributional and functional form assumptions. We assume that the error terms in both equations are normally distributed, with mean zero and unitary variance, as in standard probit models. We do not need to make assumptions about their correlation.

The estimation samples are, however, different. The likelihood of starting an apprenticeship is estimated on all graduates, whereas the estimation of the probability of passing the bar exam is restricted to the sample of those who actually sit the exam.Unfortunately, we do not have direct information about whether someone actually sits the exam, but we can approximate it quite precisely with those who did an apprenticeship. This is the only group of individuals who can take the exam and, given the length of the apprenticeship, it is unlikely that someone in this group does not take it.

Eventually, we adopt the following specifications for the probability of doing an apprenticeship and the probability of passing the bar exam:

$$\begin{aligned} P(T_i = 1|Z_i) &= \Phi\{\theta_0^T + \theta_1^T S_i + \theta_2^T N_i + \theta_3^T (S_i \times N_i) + \\ &\quad \theta_4^T A_i + \theta_5^T G_i + \theta_6^T X_i + \theta_7^T W_i\} \quad (1.7) \\ P(L_i = 1|T_i = 1, Z_i) &= \Phi\{\theta_0^L + \theta_1^L S_i + \theta_2^L N_i + \theta_3^L (S_i \times N_i) + \\ &\quad \theta_4^L A_i + \theta_5^L G_i + \theta_6^L X_i + \theta_7^L \delta_{ir}\} \quad (1.8) \end{aligned}$$

T_i is a dummy equal to one for all those graduates who report having started or completed an apprenticeship in one of the post-graduation surveys. L_i is a dummy equal to 1 if individual i eventually appears in one of the lawyers' registers within 5 years since graduation. We use Z_i to indicate the full set of explanatory variables, namely $\{A_i, S_i, N_i, G_i, W_i, X_i, \delta_{ir}\}$, where δ_{ir} indicates the fixed effect for the randomly assigned district r marking written exams in the year in which i was expected to take sit it, which we set at three years after graduation. Following conventional notation, $\Phi(\cdot)$ is the cumulative density of the standard normal distribution. Results are reported in Table 1.5.

Table 1.5: Probabilities of apprenticeship and exam

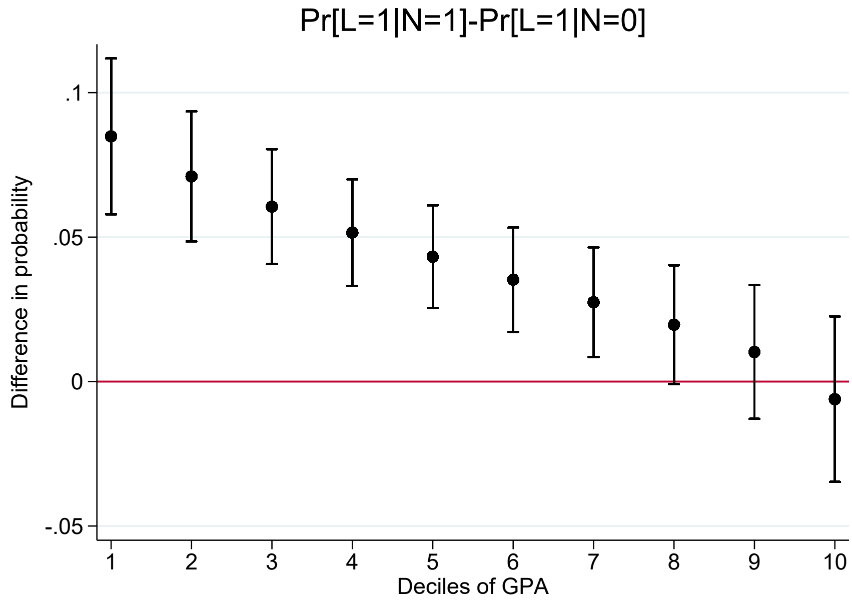
	Probability of	
	doing an apprenticeship	passing the exam
	$P(T_i = 1 \mid Z_{ir})$	$P(E_i = 1 \mid T_i = 1, Z_{ir})$
GPA[a]	0.026***	0.014**
	(0.004)	(0.006)
1=connections[b]	0.007	0.045***
	(0.005)	(0.009)
GPA × [1=connections]	-0.008*	-0.027***
	(0.004)	(0.007)
High school grade[c]	-0.014***	-0.006
	(0.002)	(0.004)
1=female	0.012***	-0.033***
	(0.004)	(0.007)
1=graduate parent[d]	0.005	-0.006
	(0.005)	(0.007)
1=parent(s) in high-ranked occup.[e]	0.016***	-
	(0.004)	
grading district FE[f]	No	Yes
Observations	24,256	21,394
Mean of dep. variable	0.882	0.544

[a] Average grade of all exams in law school. Standardised within each university. [b] 1=some connections; 0=no connections. [c] Standardised over the sample. [d] At least one parent with university degree. [e] At least one parent employed as professional, entrepreneur or manager. [f] Fixed effects for the district of exam correction.
All regressions include fixed effects for university, district, year of graduation, log size of district and log name frequency in district. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

We find that GPA matters for both passing the exam and deciding to undertake the apprenticeship period, while connections matter only for the former. The interaction of these two terms is negative and significant only in the probability of passing the bar exam. This suggests that GPA is more important for unconnected candidates than connected ones. To further investigate this important issue and get a sense of the magnitudes, Figure 1.5 shows the differences in the predicted probabilities of passing the bar exam for connected and unconnected candidates by deciles of the distribution of GPA.

Connected candidates are systematically more likely to pass the exam, especially at low levels of GPA where the difference is of almost 10 percentage points. Interestingly, GPA matters the most for unconnected candidates and very little for connected ones. As GPA increases, the gap between unconnected and connected

Figure 1.5: Difference in predicted pass rates by GPA and connections



Note: Predictions based on estimates Table 1.5, column 2. The vertical bars represent 95% confidence intervals.

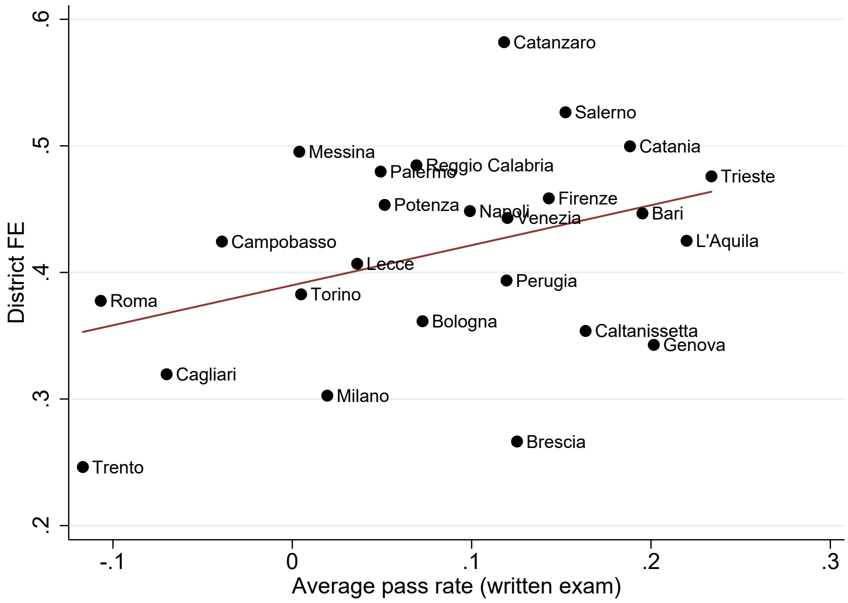
candidates narrows until it finally disappears from the eighth decile onward.

In Figure 1.6 we provide evidence of the important role of the marking commissions in the selection process. The Figure reports the fixed effects for the district of exam correction, as estimated from equation (1.8), against the pass rates at the written exam of the corresponding districts. For example, the fixed effect of the district of Milan is associated with the average pass rates at the written exams of the districts that were randomly matched with Milan over the period of our data.

Reassuringly, larger fixed effects are associated with higher average pass rates, supporting our intuition that they capture the heterogeneity in grading standards already documented in Figure 1.1. The differences are non-negligible: the predicted probability to pass the exam when the written test is graded by Trento – the district with the lowest estimated fixed effect – is 13 percentage points lower than when Trieste – the district with the highest estimated fixed effect – is grading.

These results corroborate the use of the grading district fixed effects for the identification of the earnings model of the next section.

Figure 1.6: District fixed effects and pass rates at the written exam



Note: Fixed effects for the district of exam correction from Table 1.5, column 2.

1.5.3 Earnings

We now present results from the estimation of equation (1.1), once again assuming linear functional forms:

$$y_i^L = \alpha_0^L + \alpha_1^L S_i + \alpha_2^L N_i + \alpha_3^L (S_i \times N_i) + \alpha_4^L A_i + \alpha_5^L G_i + \alpha_6^L X_i + \nu_i^L \quad (1.9)$$

$$y_i^0 = \alpha_0^0 + \alpha_1^0 S_i + \alpha_2^0 N_i + \alpha_3^0 (S_i \times N_i) + \alpha_4^0 A_i + \alpha_5^0 G_i + \alpha_6^0 X_i + \nu_i^0 \quad (1.10)$$

where y_i^J is the log of the monthly earnings that individual i self-reported in the 5-year post-graduation survey. To avoid dropping individuals with zero earnings, which is not uncommon for self-employed professionals, we simply add one to all records.[33]

Equations (1.9) and (1.10) are estimated on different samples, lawyers and non-lawyers respectively. As individuals endogenously sort into these two groups, we estimate the equations jointly using a switching regression model, where the selection equation is the combined probability of both doing an apprenticeship

[33] y_i^L and y_i^0 are meant to measure the monetary returns from alternative occupational choices and we do not model explicitly the process of finding employment.

and passing the bar exam:

$$\begin{aligned} P(L_i = 1|Z_{ir}) &= \Phi\left[\theta_0 + \theta_1 S_i + \theta_2 N_i + \theta_3 (S_i \times N_i) + \right. \\ &\left. \theta_4 A_i + \theta_5 G_i + \theta_6 W_i + \theta_7 R_r + \theta_8 X_i\right] \end{aligned} \tag{1.11}$$

The dummies for the randomly assigned grading districts are the exclusion restrictions and guarantee that identification does not rest exclusively on the arbitrarily chosen distributional assumptions. Under our basic set of assumptions, also W_i, our indicator of socioeconomic background based on parental occupation, is an exclusion restriction. We acknowledge that the exclusion assumption is more questionable for this variable than for r. We experimented with a version of the model that does not use W_i as an exclusion restriction and the results are very similar to those reported here.[34]

The estimation proceeds in two steps. In the first step, we estimate the selection equation (1.11). Then, we construct Heckman-style selection correction terms for lawyers and non-lawyers, respectively.[35] In the second step, we estimate equations (1.9) and (1.10) with OLS, each augmented with the respective selection term. Standard errors are obtained by bootstrapping.[36] The switching regression model is necessary to estimate the parameters of equations (1.9) and (1.10) avoiding bias induced by endogenous selection into occupations. The exclusion restrictions, particularly the randomly assigned marking commissions, guarantee the identification of the selection terms. Controlling for selection is important because otherwise, we would not be able to say whether connections affect earnings directly or whether the estimated coefficients merely reflect the fact that connected graduates are more likely to sort into the legal profession where earnings are higher. The same argument holds for all explanatory variables, including GPA.

Results are reported in Table 1.6. We find that a higher GPA commands higher earnings in all occupations, but more so in the legal profession. This is perfectly consistent with the idea that GPA in law school captures abilities that are more valuable in the legal profession than elsewhere. Notice also that high school grades are more important for non-legal earnings, presumably because this variable captures the returns to a broader set of skills in the absence of a measure of occupation-specific ability. Consistent with a large body of empirical evidence, we also find sizeable gender gaps in earnings.

[34] Results are available upon request.

[35] The correction terms are constructed as the ratios of the normal densities and the normal cumulative densities computed at the linear prediction and minus the linear prediction of equation (1.11), respectively for selection into the legal and non-legal occupation.

[36] See Maddala (1983) for a thorough discussion of switching regression models.

Table 1.6: Lawyer and non-lawyer earnings

	Lawyer earnings y_i^L	Non-lawyer earnings y_i^0	Selection $P(L_i = 1 \| Z_{ir})$
GPA[a]	0.237**	0.133***	0.073***
	(0.039)	(0.046)	(0.015)
1=connections[b]	0.173***	-0.096	0.122***
	(0.057)	(0.084)	(0.024)
GPA × [1=connections]	-0.122**	0.026	-0.081***
	(0.045)	(0.056)	(0.018)
High school grade[c]	0.063***	0.088***	-0.032***
	(0.024)	(0.031)	(0.010)
1=female	-0.627***	-0.658***	-0.067***
	(0.044)	(0.057)	(0.018)
1=graduate parent[d]	0.045	0.031	-0.030
	(0.043)	(0.054)	(0.020)
Exclusion restrictions:			
1=parent(s) in high-ranked occup.[e]	-	-	0.081***
			(0.019)
Grading district FE[f]	No	No	Yes
Chi-sq. of exclusion restrictions	-	-	58.17
Prob > Chi-sq.	-	-	0.000
Observations	24,260	24,260	24,260
Mean of dep. variable	5.960	5.691	0.544

[a] Average grade of all exams in law school. Standardised within each university. [b] 1=some connections; 0=no connections. [c] Standardised over the sample. [d] At least one parent with university degree. [e] At least one parent employed as professional, entrepreneur or manager. [f] Fixed effects for the district of exam correction.
All regressions include fixed effects for university, district, year of graduation, log size of district and log name frequency in district. Bootstrapped standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

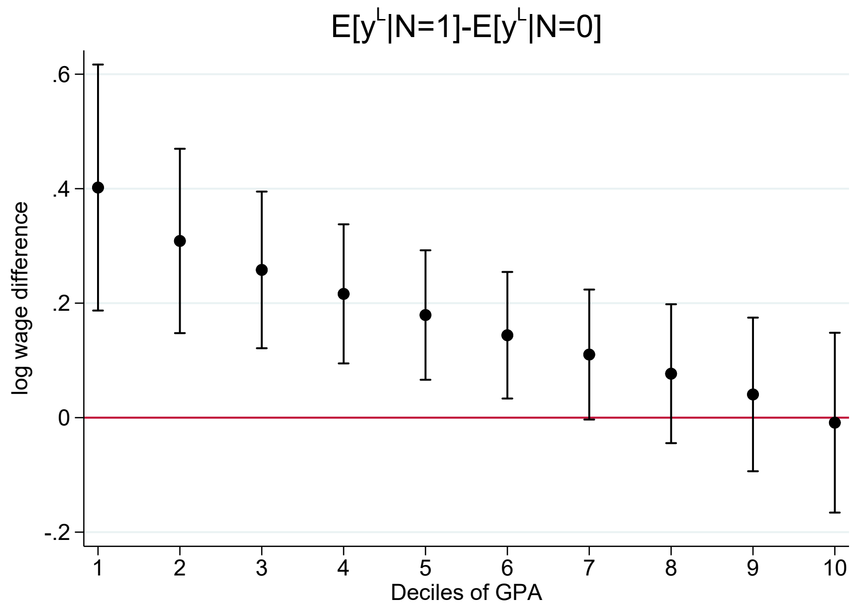
Having relatives in the profession is also associated with higher earnings, but only for those working as licensed lawyers. In addition, the interaction of GPA and connectedness is negative for lawyers (and non-significant for non-lawyers).

The exclusion restrictions also work as expected. The children of parents employed in high-ranked occupations are more likely to become lawyers because they are in a better position to sustain the costs of the long preparation. The randomly assigned grading district also matters substantially for the probability of passing the exam.

The test of the joint significance of both the grading district dummies and parents' occupation solidly rejects the null hypothesis, as reported in Table 1.6. We obtain a similar result also when testing the significance of the set of grading district dummies alone (the Chi-squared statistics is equal to 40.44, with a p-value of 0.026).

To get a better sense of the magnitudes of the effects implied by the estimates in Table 1.6, Figure 1.7 shows the predicted difference in log earnings in the legal profession between connected and unconnected lawyers along the distribution of GPA. In the bottom decile of the distribution, connected lawyers earn 40% more than their unconnected colleagues and it is only towards the very top of the distribution that this difference becomes statistically insignificant.

Figure 1.7: Predicted (log) wage differences between connected and unconnected lawyers by GPA



Notes: Predictions based on the estimates of Table 1.6, columns 1 and 2. The vertical bars represent 95% confidence intervals.

One possible interpretation of this effect is that unconnected lawyers, especially at the beginning of their careers, find it difficult to attract clients. Having relatives who can share their portfolios of clients might represent a significant advantage. This interpretation is also consistent with the strict regulations concerning professional practice. In Italy, like in many other countries in continental Europe, professional associations impose codes of conduct that regulate commercial practices, among other things. For example, it is often prohibited to approach clients who are already served by another professional and, until recently, commercial advertising was considered to be contrary to the "dignity of the profession". In addition, the code of conduct indicates price floors.[37] Being unable to lower prices and advertise their services, young lawyers find it extremely difficult to attract

[37]Some of these regulations were recently reformed in Italy, but they remain strongly present in the daily practice of the profession Orsini and Pellizzari (2012).

clients.

An alternative interpretation of this effect is that connected lawyers, especially those with low GPA, have higher occupation-specific human capital along dimensions that are not captured by GPA. In Section 1.5.4, we present various pieces of evidence that seem to contradict such an alternative explanation.

1.5.3.1 Working with relatives

In this Section, we present additional evidence showing that the effect of connections on earnings in the legal profession is substantially larger when young lawyers work in the same law firms as their connections. We do not formally incorporate the process of selecting into law firms into our model of Section 1.4 because its identification would require an additional exclusion restriction, which we do not have. Hence, these results need to be interpreted with caution.

Using our proxy of law firms based on professional coordinates (see Section 1.3 for details), we find that about 7% of the lawyers in our data work with relatives. We augment equation (1.9) with a term indicating whether the young lawyer works in the same law firm with someone holding her/his same surname. Let F_i be such an indicator. We then estimate the following equation:

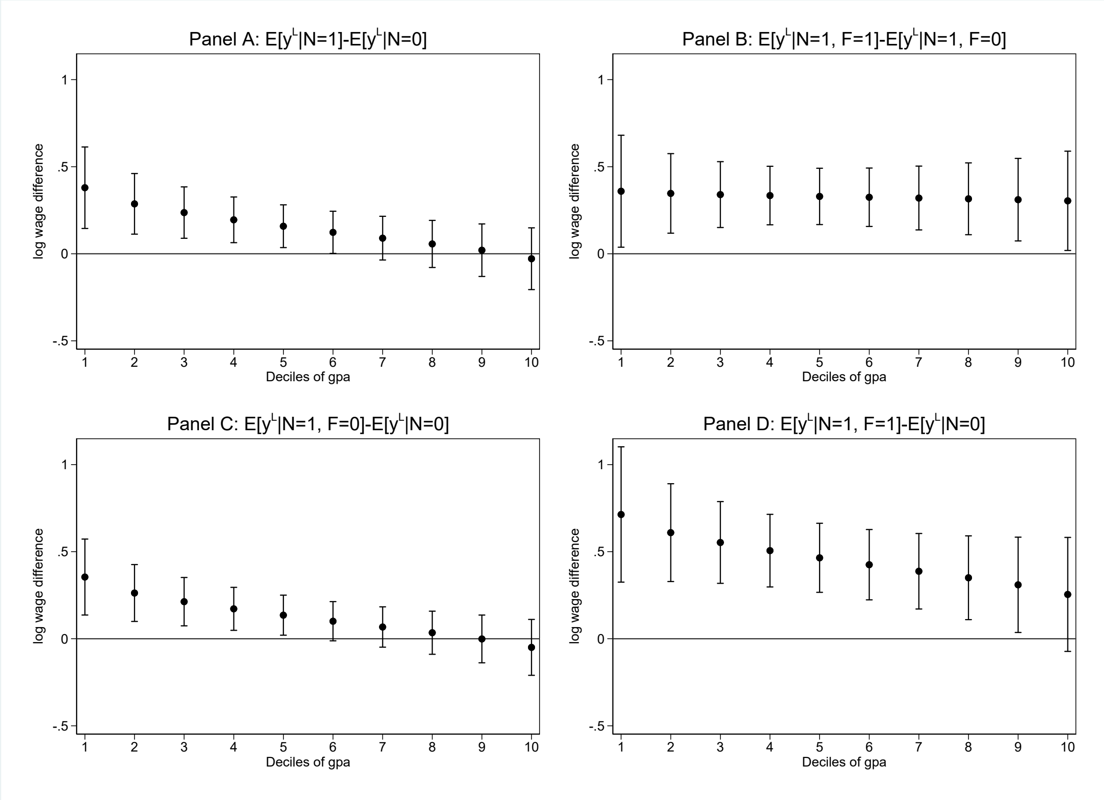
$$y_i^L = \alpha_0^L + \alpha_1^L S_i + \alpha_2^L N_i + \alpha_3^L (S_i \times N_i) + \alpha_4^L A_i + \alpha_5^L G_i + \alpha_6^L X_i + \alpha_7^L F_i + \alpha_8^L (F_i \times S_i) + \nu_i^L \tag{1.12}$$

For brevity, we only report results graphically. Figure 1.8 follows the same logic as Figure 1.7, but it extends the comparison to lawyers working in the same firms as their connections or in others. For completeness, Panel A replicates under the specification of equation (1.12), the same analysis of Figure 1.7, namely the comparison of the average earnings of connected and unconnected individuals, regardless of which firm they work in. The three subsequent panels decompose this earnings gap by both connectedness status and firm type.

Panel B of Figure 1.8 focuses exclusively on connected individuals and compares the average earnings of those working with relatives with those who do not. Throughout the distribution of GPA, young lawyers working in the same law firm as a relative earn around 4% more than colleagues who, despite having family connections with the profession, do not work with them in the same firm. As both groups are connected and thus potentially benefit from the same intergenerational transmission of occupation-specific human capital, the gap can be rationalized through preferential access to a stock of clients by young lawyers who work with their relatives.

Panel C compares connected lawyers who do not work with their relatives against unconnected colleagues. The resulting figure is very similar to Panel A,

Figure 1.8: Predicted (log) wage differences by GPA, connectedness and firm type



Notes: In Panel A, we show the baseline result (wage differences between connected vs. unconnected lawyers). In Panel B, we show the wage differences of connected lawyers working with relatives with those who do not. Panel C reports the wage differences of connected lawyers not working with their relatives against unconnected lawyers. Finally, Panel D reports the wage differences between connected lawyers working with relatives and unconnected ones. The vertical bars represent 95% confidence intervals.

with a difference of 2.5-3 percentage points in the lowest deciles that vanishes approximately above the median. Even though we cannot rule out that there is no transmission of occupation-specific human capital among relatives, it is unlikely that it only occurs for lawyers whose GPA is below the median. As discussed below, these findings suggest that family connections increase wages by facilitating access to clients, rather than improving young lawyers' preparedness.

Finally, Panel D compares connected individuals working with relatives and unconnected individuals and shows a large 6.5 percentage points differential in average earnings at low-GPA deciles, which shrinks modestly as GPA increases. Only in the tenth decile, this difference in earnings becomes statistically insignificant (at the 95% level).

Overall, these results show that when connected lawyers work in the same firm as their family ties their earnings advantage over unconnected colleagues increases substantially and remains significant across almost the entire distribution of GPA.

1.5.4 Professional ability, GPA and family connections

Some of our most important findings can be subject to alternative interpretations. For example, the evidence in Figure 1.5 shows that connected candidates are more likely to pass the entry exam, especially at low levels of GPA. This result could be explained by some form of nepotism. Established senior lawyers may lobby or put pressure on colleagues who sit in the exam commissions to facilitate entry into the profession by young members of their families who might otherwise be unlikely to pass the exam. Alternatively, connected candidates might be more likely to pass the entry exam compared to their unconnected colleagues because they have higher occupation-specific human capital that they have accumulated thanks to their family connections. By internalising the possibility to learn the trade in the family, these young connected professionals may choose to exert little effort during law school, thus explaining their low GPA.

The finding in Figure 1.7 can also be seen under alternative interpretations. The figure shows that connected lawyers earn more than their unconnected colleagues, especially at low levels of GPA. One interpretation of this result is that being connected to a senior professional helps access clients and generate revenues, especially for the least able entrants in the market for legal services. Young lawyers with high GPA can find clients even without the help of their connections, so it does not really matter whether they have any. An alternative explanation is, once again, that young lawyers with lower GPA have higher professional ability and that is the simple reason why they earn more than others.

In this section, we present four pieces of evidence that help us assess the relative validity of these alternative explanations. In addition, in Appendix A we develop a simple model of professional human capital formation that allows accumulation to take place in multiple settings, such as in law school and during the apprenticeship or in the family. This model informs us about the features of the human capital formation process that could rationalise the idea that lawyers with lower GPA have higher professional ability. Without going through much detail, this theoretical investigation suggests that, as long as the human capital accumulated in the various settings is sufficiently complementary, GPA would positively correlate with professional ability.

The first piece of empirical evidence that we present here looks at the size of the law firms where the young lawyers in our data eventually work. If one is willing to assume that the larger firms are also better and more productive, then one would expect the best professionals to work there. We are aware that firm size is not a perfect measure of firm quality but, consistent with a large literature, we find a substantial firm-size premium in earnings in our data.[38]

Figure 1.9: Average law-firm size by connectedness and GPA

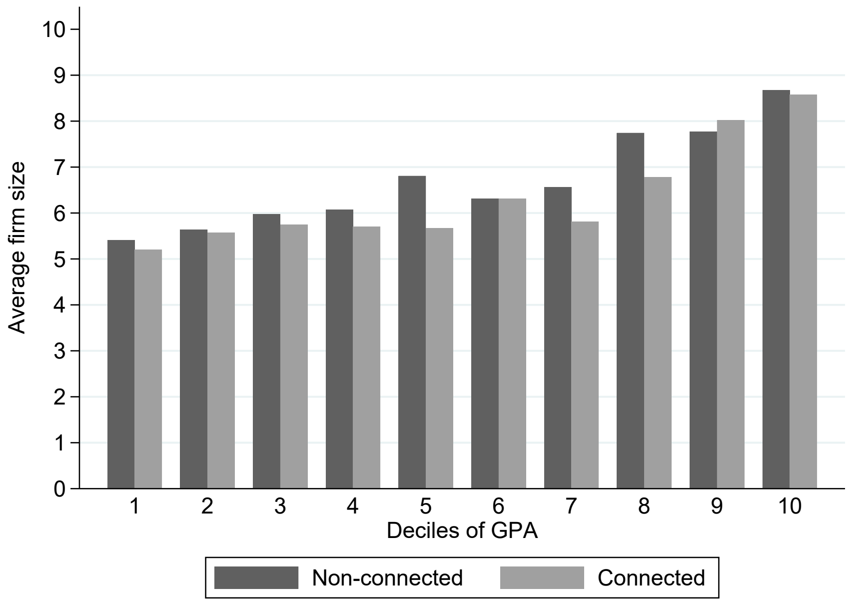


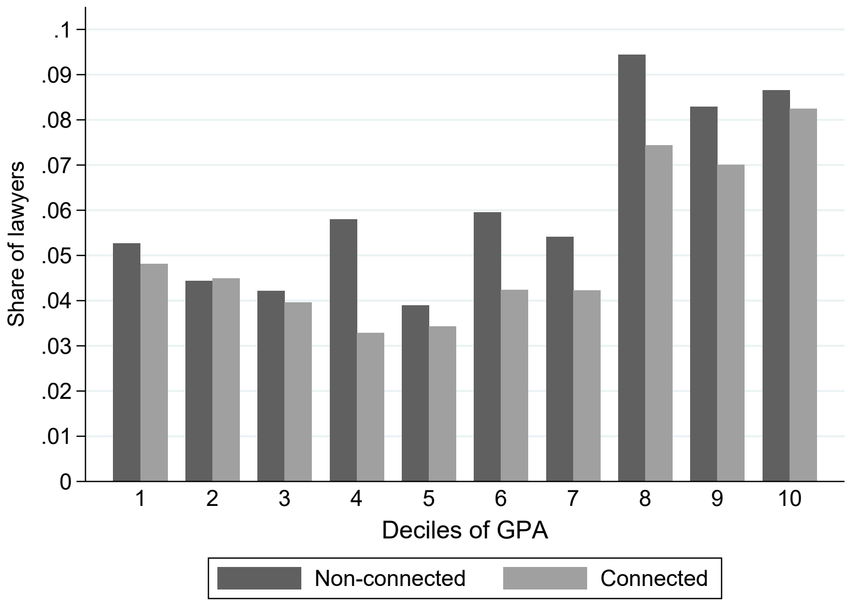
Figure 1.9 shows the average firm size of connected and unconnected lawyers by deciles of GPA, and it clearly indicates that professionals with higher GPA tend to work in larger firms, regardless of connections. This result seems at odds with the idea that connected lawyers with low GPA have high professional ability. If this

[38]In unreported regressions, we find that the earnings of those who work in law firms with more than 8 colleagues (top 25%) earn 75% more than those who are self-employed. Results are available from the authors upon request.

were the case, one would expect them to work in larger firms than unconnected colleagues, both those with low and high GPA. This is clearly not the evidence shown in Figure 1.9.

Second, we look at geographical movers. Figure 1.10 shows the share of lawyers in our data who move to the largest cities of the country: Milan, Naples, and Rome. These large cities are also the most remunerative markets for legal services and one would expect only the most skilled professionals to be successful there.[39] We define movers as young lawyers who are observed in the professional registers of one of these big cities and who went to law school in a different district.

Figure 1.10: Shares of lawyers moving to the largest cities by connectedness and GPA



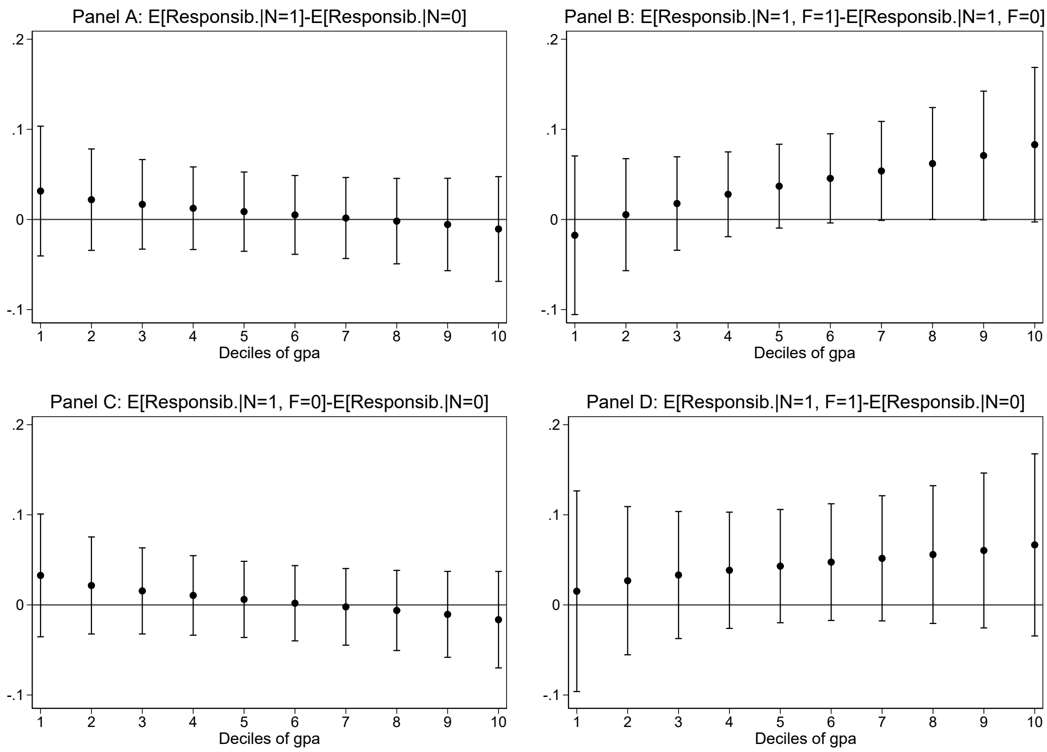
Note: the largest cities are Milan, Naples and Rome.

Despite some nonlinearities, results show that the lawyers with higher GPA are more likely to move to large cities, with small differences between the connected and the unconnected. If anything, unconnected lawyers seem to be slightly more likely to move, especially at the top of the GPA distribution. Similar to the evidence in Figure 1.9, this finding also counters the idea that connected lawyers with low GPA are highly skilled. If this were the case, one would expect them to be more likely to move to the most remunerative markets, which is not what Figure 1.10 indicates. Notice that even though both Figure 1.9 and Figure 1.10 are based on simple descriptive statistics, the results do not change when conditioning on the set of controls that we use in equation (1.9).

[39]These results are robust to the definition of most remunerative city (available upon request).

Next, we exploit the information available in the post-graduation surveys on work activities. Particularly useful for our purposes is the question in which individuals are asked whether they contribute to the definition of the strategies of the firm they work for. Presumably, these responsibilities are more likely to be assigned to more skilled individuals. We construct a dummy indicator for whether respondents report contributing to the definition of the strategies of their firms and Figure 1.11 presents results from the estimation of a switching regression model similar to equations (1.9)-(1.10) but with our indicator of work responsibility as a dependent variable.

Figure 1.11: Predicted differences in work responsibilities by GPA, connectedness and firm type



Notes: Predictions based on the estimates of equations (1.9), (1.10) and (1.11), with the probability of carrying out strategic work responsibilities as the dependent variable. The vertical bars represent 95% confidence intervals.

Like Figure 1.8, Panel A compares connected and unconnected individuals, regardless of which firm they work in. Results show not only no significant differences between connected and unconnected lawyers, but also no significant pattern by decile of GPA. Thus, this evidence seems inconsistent with the idea that connected lawyers have higher occupation-specific ability, especially at low levels of GPA.

Panel B, again, focuses exclusively on connected individuals and compares those who work with relatives with those who do not. Interestingly, the former tend to contribute more to the definition of the strategies of the firm, the higher their GPA. Assuming that one's family connections know one's ability more accurately than others, this comparison seems to rule out the possibility that the findings in Panel A might be due to some form of informational asymmetry.

Panel C compares connected lawyers who do not work with their relatives with their unconnected colleagues. The resulting figure is very similar to Panel A. Finally, Panel D shows that connected lawyers that work in the same firm as their family ties do not exhibit any significant differences with respect to unconnected lawyers. Additionally, no significant pattern emerges along the distribution of GPA.

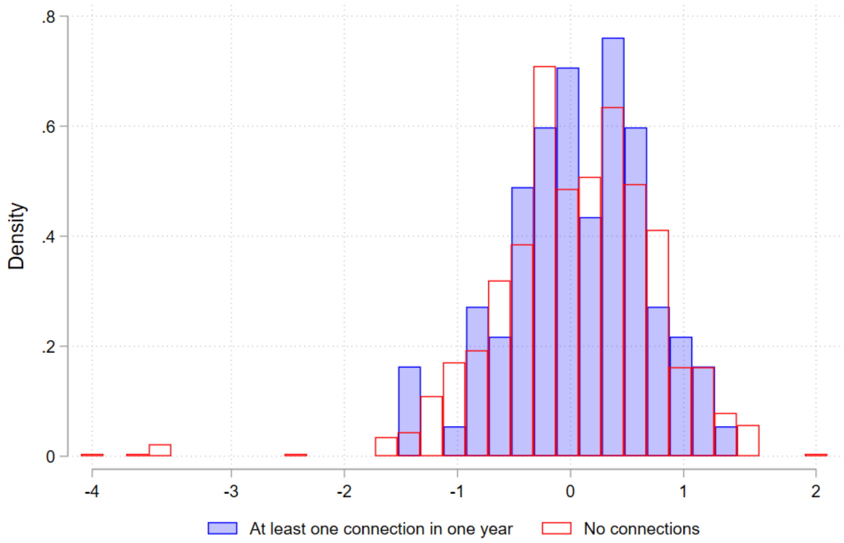
Our final piece of evidence investigates the possibility that connected lawyers choose to exert little effort in law school – thus obtaining a low GPA – and recoup the lost human capital later on through apprenticeships in high-quality law firms, which they access thanks to their connections. In this case, GPA in law school might even be negatively correlated with professional ability at the end of the apprenticeship.

To examine the plausibility of this scenario, we explore whether connected individuals do their apprenticeships at law firms of particularly high quality. We do this by exploiting an auxiliary dataset covering the population of all lawyers (both junior and senior) and apprentices of one Italian region, Veneto. In this dataset, we observe both incomes and the law firm in which the lawyers and the apprentices are employed. We also observe surnames and other basic demographic information.[40] We use this dataset to estimate law firm fixed effects from a regression of lawyers' earnings.[41] We interpret these fixed effects as measures of the quality of the law firms and we match them to the apprentices doing their training in these law firms. Notice that we could not perform this exercise on our main dataset because we do not observe apprentices in their law firms. Using the Veneto data, We can then check whether connected apprentices sort into firms with higher fixed effects compared to similar unconnected apprentices.

[40]Veneto registers, excluding that of Venice, were collected in the context of a previous project focusing on this region (see Pellizzari and Pica (2010)). Contrary to the data used in this paper, in Veneto we only observe lawyers and apprentices and we have no information on law graduates who choose a different occupation. Hence, we cannot study selection into the legal profession with that dataset.

[41]To estimate firms fixed effects, we restrict the analysis to licensed lawyers (i.e., we exclude apprentices, who should have zero earnings from legal practice) and regress (log) income on a gender dummy, a connection dummy, experience dummies, age dummies, region of birth dummies, the total number of surnames within each register/year, the total number of surnames within each province, year dummies and firms fixed effects. In this ancillary dataset, we do not have information on individual GPA, which implies that the firm fixed effect also captures the ability of the lawyers employed in the firm. We do not find this problematic for this specific exercise.

Figure 1.12: Distribution of law-firm fixed effects by connected and unconnected apprentices



Note: Firm fixed effects are produced from a regression of (log) income on firm dummies for the population of lawyers of the Veneto region. The additional controls included in the regressions are a gender dummy, a connection dummy, experience dummies, age dummies, region of birth dummies, total number of surnames within register/year, total number of surnames within the province and year dummies. Apprentices are excluded when calculating the fixed effects; lawyers from Venice are not part of the dataset (Pellizzari and Pica, 2010).

Figure 1.12 overlays the distributions of the firm fixed-effects for connected and unconnected apprentices and shows that the two overlap substantially. This result is at odds with the possibility that connected apprentices train at better firms.

Overall, the results in this section seem to be more consistent with the interpretation of our main findings based on some form of nepotism rather than human capital accumulation within professional dynasties. We acknowledge that none of the individual pieces of evidence is entirely conclusive but we believe that their collection makes it quite unlikely that our results can be generated by young lawyers with low GPA having high professional ability. This interpretation is also consistent with what we believe to be the most natural assumption about the process of human capital formation described in our theoretical model in Appendix A, namely a certain degree of complementarity between the human capital accumulated in different settings.

1.6 Simulations

Using the estimates of our model, we can perform counterfactual exercises. We are particularly interested in understanding the role of connections in the selection process into the legal profession. In our model of Section 1.4 there exist two potential channels through which family ties could influence the process of occupational choice. First, connected candidates are apparently facilitated in passing the bar exam. Figure 1.5 suggests that the effect of connections on the probability of passing the exam is stronger at lower levels of GPA, thus potentially generating negative selection or, at least, mitigating positive selection along this dimension. Second, connected individuals earn higher earnings than other colleagues and, once again, the effect is stronger at the bottom of the distribution of GPA. To the extent that individuals are forward-looking, we expect also this second channel to generate negative selection on academic ability.

Entering the legal profession is the combined outcome of two events. First, one needs to do the compulsory apprenticeship and, then, one needs to pass the bar exam. Our model describes these events in equations (1.3) and (1.4), respectively, and in Section 1.5 we have produced estimates of their probabilities. However, in order to separately identify the different channels through which family ties affect the process, we need to modify the way we estimate the choice of an apprenticeship. In Section (1.5) we estimated it as described in equation (1.7), which does not allow disentangling the role of connectedness on earnings and the probability of passing the bar exam.

Hence, we go back to the definition of the probability of doing an apprenticeship presented in the theoretical Section (1.4), equation (1.4). First, we use the estimates of equations (1.3), (1.9) and (1.10) to compute the expected earnings premium in the legal profession (conditional on doing the apprenticeship):

$$\widehat{E}\left[\Delta y_i | Z_{ir}\right] = \widehat{y}^L(A_i, S_i, N_i.G_i, X_i) - \widehat{y}^0(A_i, S_i, N_i, G_i, X_i) \tag{1.13}$$

Then, we re-estimate the probability of the apprenticeship directly from its theoretical definition in equation (1.4):

$$\begin{aligned} \widehat{P}(T_i = 1 | Z_{ir}) &= P\left[v_i < \widehat{P}(L_i = 1 | T_i = 1, Z_{ir})\widehat{E}[\Delta y_i | Z_{ir}] - \widehat{\theta}_6^T W_i\right] = \\ & \Phi\left[\widehat{P}(L_i = 1 | T_i = 1, Z_{ir})\widehat{E}[\Delta y_i | Z_{ir}] - \widehat{\theta}_6^T W_i\right] \end{aligned} \tag{1.14}$$

where $\Phi(\cdot)$ is the cumulative density of the standard normal distribution. Notice that we had already assumed normality of v_i in Section 1.5, so there are no additional assumptions in equation (1.14).[42]

Finally, we estimate the probability of being a lawyer as follows:

$$\begin{aligned}\widehat{P}(L_i = 1|Z_{ir}) &= \widehat{P}(L_i = 1|T_i = 1, Z_{ir})\widehat{P}(T_i = 1|Z_{ir}) \\ &= \widehat{P}(L_i = 1|T_i = 1, Z_{ir}) \\ &\quad \Phi\left[\widehat{P}(L_i = 1|T_i = 1, Z_{ir})\widehat{E}[\Delta y_i|Z_{ir}] - \widehat{\theta}_6^T W_i\right]\end{aligned} \tag{1.15}$$

To carry out the counterfactual analysis, we compare the predicted outcomes obtained through equations (1.13)-(1.15) estimated with the dataset's measure of connections with the predictions from the hypothetical scenario in which there are no connections, $N_i = 0$ for all i, where the population of law school graduates is kept constant. Both scenarios use the coefficients obtained from estimating equations (1.13)-(1.15) with the full dataset and differ only because the predicted outcomes in the latter are estimated without connections. The main underlying assumption that allows this comparison is that selection into law school is independent of connections. In other words, eliminating connections does not change the composition of the population of law school graduates. Using similar data to investigate the intergenerational transmission of liberal professions, Aina and Nicoletti (2018) find that parents who are liberal professionals do not affect children's choice of major at university while affecting the choice of liberal occupation. Hence, we expect low levels of outmigration from law school due to a lack of connections.

Panel A of Figure 1.13 compares the average estimated probability of being a lawyer from equation (1.15) with the share of lawyers in the raw data, breaking down the results by deciles of the distribution of GPA. Although the model predicts slightly lower incidence of lawyers at the bottom of the distribution and slightly higher at the top, the overall fit is quite good and we can replicate the small degree of positive selection that is observed in the data. In fact, the confidence intervals of the model predictions overlap with those of the data for all deciles.[43]

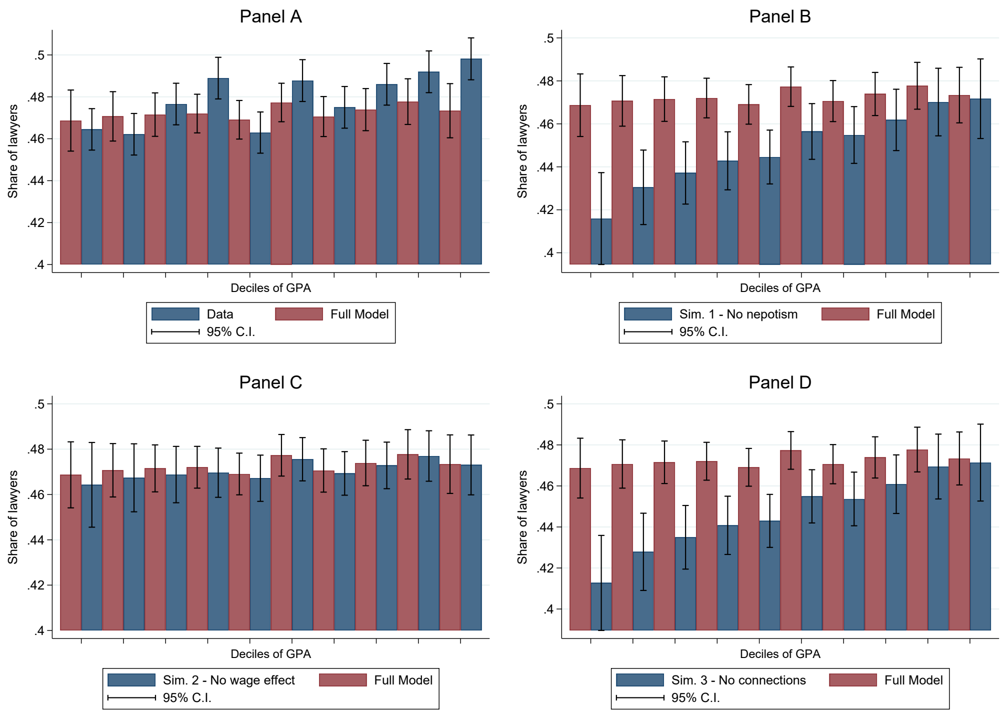
The following panels replicate the simulations of the selection probabilities under different scenarios and compare results with the predictions of the original model, i.e. those reported in the first panel.

Panel B shows results produced by eliminating the influence of family connections from the probability of passing the bar exam, but not from the earnings

[42] To improve the accuracy of our predictions, we actually estimate $\Phi\left[\widehat{P}(L_i = 1|T_i = 1, Z_{ir})\widehat{E}[\Delta y_i|Z_{ir}] - \widehat{\theta}_6^T W_i\right]$ as a probit model with $\widehat{P}(L_i = 1|T_i = 1, Z_{ir})\widehat{E}[\Delta y_i|Z_{ir}]$ and W_i as explanatory variables.

[43] The confidence intervals have been generated replicating the calculations on both the raw data and the model over 1,000 bootstrapped samples. We also checked that the model predicts well other relevant outcomes, such as earnings or the probability of doing the apprenticeship. Results are available upon request.

Figure 1.13: Counterfactual simulations exercises



Notes: The Figure reports the results of the counterfactual simulation exercises described in Section 1.6. Panel A reports the baseline results of the empirical model. Panel B reports the results of the simulations with no connections at the exam stage. Panel C reports the results of the simulations with no connections at the earnings stage. Panel D reports the results of the simulations with no connections at any stage.
The confidence intervals have been generated replicating the calculations on both the raw data and the model over 1.000 bootstrapped samples.

process. More specifically, we simulate the probability of being a lawyer as:

$$\begin{aligned} \widehat{P}_B(L_i = 1|Z'_{ir}) &= \widehat{P}(L_i = 1|T_i = 1, Z'_{ir}, N_i = 0) \\ & \Phi\left[\widehat{P}(L_i = 1|T_i = 1, Z'_{ir}, N_i = 0)\widehat{E}[\Delta y_i|Z_{ir}] - \widehat{\theta}_6^T W_i\right] \end{aligned} \quad (1.16)$$

where Z'_{ir} is the set of all explanatory variables of the model, excluding the dummy indicator of connected individuals N_i, $Z'_{ir} = \{A_i, S_i, G_i, W_i, X_i, R_r\}$. Results show that, when family connections do not influence the results of the entry exam, a substantial degree of positive selection on GPA emerges, especially due to fewer individuals with low GPA entering the profession. The simulation shows that, compared to the original model, the predicted share of lawyers declines by over 4 percentage points (from 0.45 to 0.41) in the lowest decile of GPA, whereas it increases by one percentage point at the top.

In Panel C, we repeat the simulation exercise, but this time we eliminate the effect of family connections from the earnings process and we maintain it in the exam:

$$
\begin{aligned}
\widehat{P}_C(L_i = 1|Z'_{ir}) &= \widehat{P}(L_i = 1|T_i = 1, Z_{ir}) \\
& \quad \Phi\left[\widehat{P}(L_i = 1|T_i = 1, Z_{ir})\widehat{E}[\Delta y_i|Z'_{ir}, N_i = 0] - \widehat{\theta}_6^T W_i\right]
\end{aligned}
\tag{1.17}
$$

Contrary to the previous analysis, the predicted shares of lawyers by deciles are now similar to the original model, suggesting that the effect of family connections on earnings has limited influence on the selection process.

Finally, in Panel D we consider a scenario in which family connections have no influence, neither on earnings nor on the exam:

$$
\begin{aligned}
\widehat{P}_D(L_i = 1|Z'_{ir}) &= \widehat{P}(L_i = 1|T_i = 1, Z'_{ir}, N_i = 0) \\
& \quad \Phi\left[\widehat{P}(L_i = 1|T_i = 1, Z'_{ir}, , N_i = 0)\widehat{E}[\Delta y_i|Z'_{ir}, N_i = 0] - \widehat{\theta}_6^T W_i\right]
\end{aligned}
\tag{1.18}
$$

Consistently with the previous simulations, we find that positive selection is now much stronger than in the original model and the results are numerically very similar to those in Panel B.

Taken together, these simulations point to the fact that without connections there would be significantly fewer lawyers with low GPA, while high-GPA lawyers would not be penalized. In addition, eliminating connections would also slightly reduce the overall pass rate, hence the overall number of licensed lawyers. The average simulated pass rate declines from 47.2% in the model with connections to 44.8% when connections are completely eliminated.[44] Overall, without connections there would be fewer lawyers on the market with comparatively higher GPA, and the main channel through which connections impact the probability of becoming a lawyer is through the probability of passing the exam.[45]

Given the results presented in Section 1.5.4, we believe that the most likely interpretation of these findings is related to nepotistic practices in the entry exam. Presumably, this happens due to two factors: (i) the important role of incumbent lawyers in the process and (ii) the partial anonymity of the examination (in our specific case, this is due to the oral interview). In many systems of occupational regulation around the world, especially with regard to liberal professions, either one or the other or both of these factors are present and our results could easily generalise to most of these settings.

[44] Simulations with a fixed pass rate (either overall or by district) show very similar results in terms in selection.

[45] Alternative simulation exercises in which all individuals are corrected either by the most lenient or the strictest district show that selection varies as expected.

1.7 Robustness checks

In this Section, we present several robustness checks to complement our main analysis. Section 1.7.1 investigates the implications of measurement error in family connections. In Section 1.7.2, we study how our results change when we take into account that wage growth in the legal profession might be different than in other occupations. Finally, in Section 1.7.3 we replicate our main findings using the largest possible number of observations for each equation instead of using the same sample in all of them, as we do in Section 1.5.

1.7.1 Measurement error in connectedness

We measure family connections with family names and it is rather obvious that such a measure is subject to error. The direction of the error is difficult to predict. On the one hand, we might be missing some relevant family ties who do not share the same surname as the individuals in our sample. Given that Italy adopts the relatively standard practice of giving children the surname of the father, our measure clearly misses relatives coming from the maternal arm of the family. On the other direction, there can also be individuals who share the same surname and are nevertheless not linked to each other by any kinship connection. This is especially true for frequent surnames. However, we know from previous studies that only a very small share of the population holds very frequent surnames and, for the very vast majority of cases sharing the same surname is associated with a very high probability of being related to each other via some family link (Güell et al., 2015, 2018).

Some datasets contain information on direct parent-children connections (Chetty et al., 2014; Raitano and Vona, 2021). This is not the case in our data. Notice, however, that it is difficult to say whether, for the purposes of this paper, our surname-based measure of connections is better or worse than one relying on exact parent-child links. Using surnames is subject to the mis-classification errors discussed above, but it allows capturing family ties beyond mothers and fathers, like grandparents or uncles/aunts, who might also influence one's occupational career.

Unfortunately, there is little we can do with our data to identify or reduce the error in our indicator of family connections. Hence, we take a different approach and, instead of trying to reduce mis-measurement, we artificially increase it and we look at how much more error would be necessary to make our main results go away.

For brevity, we only focus on two outcomes, namely the probability of passing the bar exam and earnings, and we re-estimate the corresponding equations using an indicator of family connections where a given share of observations are randomly re-coded.[46]

Table 1.7: Measurement error in connections and the probability of passing the bar exam

	Percentage of randomly re-assigned connections[a]					
	0%	1%	5%	10%	20%	30%
	[1]	[2]	[3]	[4]	[5]	[6]
GPA[b]	0.042**	0.036**	0.035**	0.026*	0.017	0.006
	(0.016)	(0.016)	(0.016)	(0.015)	(0.015)	(0.015)
1=connections[c]	0.128***	0.113***	0.097***	0.060***	0.020	0.010
	(0.026)	(0.025)	(0.023)	(0.022)	(0.020)	(0.019)
GPA × [1=connections]	-0.079***	-0.070***	-0.069**	-0.055**	-0.040**	-0.022
	(0.019)	(0.019)	(0.019)	(0.019)	(0.018)	(0.018)
High school grade[d]	-0.017	-0.017	-0.017	-0.018	-0.018*	-0.018*
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
1=female	-0.091***	-0.091***	-0.092***	-0.092***	-0.094***	-0.094***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
1=graduate parent[e]	-0.018	-0.017	-0.015	-0.013	-0.012	-0.012
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Observations	21,380	21,380	21,380	21,380	21,380	21,380

[a] Percentage of connected individuals who are randomly reassigned to having no connections. Each time an equal number of unconnected individuals is randomly assigned to be connected. [b] Average grade in all exams at the law school. Standardised within each university. [c] 1=some connections; 0=no connections. [d] Standardised over the sample. [e] At least one parent with a university degree.
All specifications include fixed effects for university, district, district of exam correction three years after graduation, and year of graduation. Probit coefficients are reported. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 1.7 reports the results of this exercise for the probability of passing the bar exam. For comparison purposes, the first column simply reports our main results from equation (1.8) (compare with column 2 of Table 1.5). In column 2, we replicate the same estimation, but we randomly recode 1% of the connected individuals as unconnected and we randomly take an equal number of unconnected individuals recoding them as connected.[47] The following columns perform the same exercise with higher shares of random re-classification. Of course the magnitude of the estimates changes across columns, but we find reassuring that our main results on GPA, connections and their interaction are qualitatively robust

[46]Results for the other equations confirm the findings in this section and can be obtained upon request.
[47]We also experimented with other forms of recoding, such as recoding a given share of the connected and of the unconnected, and results are consistent with what we report in this section. An advantage of our specific choice of the exercise is that the share of connected individuals remains fixed and we can associate the differences in results exclusively to mis-measurement.

and tend to disappear only when we reclassify large shares of individuals, i.e., more than 20%.

Table 1.8: Measurement error in connections and lawyers' earnings

	Percentage of randomly re-assigned connections[a]					
	0%	1%	5%	10%	20%	30%
	[1]	[2]	[3]	[4]	[5]	[6]
GPA[b]	0.237***	0.230***	0.218***	0.219***	0.182***	0.172***
	(0.039)	(0.039)	(0.038)	(0.038)	(0.036)	(0.035)
1=connections[c]	0.173***	0.174***	0.158***	0.156***	0.102**	0.043
	(0.057)	(0.056)	(0.053)	(0.049)	(0.045)	(0.042)
GPA × [1=connections]	-0.122***	-0.112**	-0.094**	-0.097**	-0.041	-0.026
	(0.045)	(0.045)	(0.044)	(0.043)	(0.042)	(0.041)
High school grade[d]	0.063***	0.063**	0.061**	0.061**	0.061**	0.061**
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
1=female	-0.627***	-0.627***	-0.627***	-0.628***	-0.629***	-0.630***
	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)
1=graduate parent[e]	0.045	0.046	0.050	0.052	0.054	0.056
	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)
Observations	24,260	24,260	24,260	24,260	24,260	24,260

[a] Percentage of connected individuals who are randomly reassigned to having no connections. Each time an equal number of unconnected individuals is randomly assigned to be connected. [b] Average grade in all exams at the law school. Standardised within each university. [c] 1=some connections; 0=no connections. [d] Standardised over the sample. [e] At least one parent with a university degree.
All specifications include fixed effects for university, district, and year of graduation. Results are obtained with a switching regression model, where the exclusion restrictions are fixed effects for the grading district and parental occupation. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

In Table 1.8 we reproduce the same exercise for the earnings equations (1.9) and (1.10), with equation (1.11) completing the switching regression model. For brevity, we only report results for earnings in the legal profession and, similarly to Table 1.7, we find that adding measurement error to our indicator of family connections only affects our coefficients of interest when we reclassify relatively large shares of individuals (above 20%).

Overall, we are reassured by the results in this section. Although we cannot exclude a priori that measurement error in connections affects the magnitude of our most important estimates, it seems unlikely that this bias is large enough to overturn their qualitative message.

1.7.2 Differential wage growth

One limitation of our data is that we observe earnings only at the very beginning of one's career. More specifically, we observe self-reported earnings via the survey carried out at five years since graduation. For rational forward-looking agents, this is not the relevant measure of earnings they should use for their occupational choices. Rather, they should consider the present discounted value of the entire stream of expected future earnings.

Unfortunately, we do not have longitudinal earnings data for lawyers and for law school graduates who entered a different occupation. Nevertheless, we have collected from external sources the average annual growth rates of earnings for these two categories of individuals, broken down by gender.

For lawyers, we obtain this information from the professional social security administration (*Cassa Forsense*).[48] For non-lawyers we compute the growth rates of earnings from the official Italian Labour Force Survey (LFS), which contains information on field of study and occupation. We pool all surveys from 2009 to 2018 and we restrict the sample to graduates from law school who are not employed in a liberal profession. With this data, we estimate cross-sectional experience profiles, separately by gender and conditional on year effects.[49]

Figure 1.14 shows the earnings profiles implied by these growth rates and, to facilitate the comparison, we normalise initial earnings to one for all four categories. Apparently, earnings in the legal profession grow more rapidly than in other professions (but for law school graduates), but they are also more concave.

To understand the implications of the differential growth rates of earnings for our simulations, Figure 1.15 reproduces the exercise by redefining expected (log) earnings as the full stream of discounted future earnings over 30 years of experience:

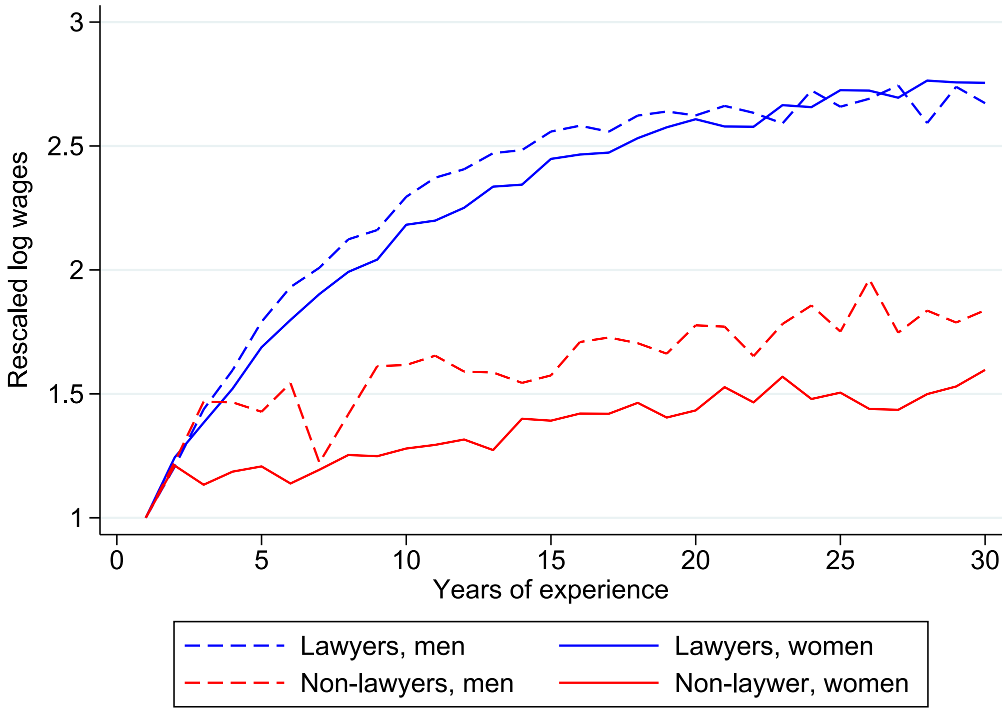
$$\widehat{E}\left[\Delta y_i|Z_{ir}\right] = \widehat{E}\left[(y_i^L + \sum_{e=1}^{30}\gamma_{ie}^L) - (y_i^0 + \sum_{e=1}^{30}\gamma_{ie}^0)|Z_{ir}\right] \tag{1.19}$$

where γ_{ie}^L and γ_{ie}^0 are the growth rates of earnings for lawyers and non-lawyers at experience e and the variation across individuals is restricted to gender. Obviously, we similarly redefine $\widehat{E}\left[\Delta y_i|Z'_{ir}, N_i = 0\right]$.

48 We thank Michele Raitano for providing us with these aggregate growth rates. The experience profiles are produced via a simple OLS regression with log earnings as a dependent variable and year dummies (data is available for 6 years: 1985, 1990, 1995, 2000, 2005 and 2008) and experience dummies (one per each year of experience) as explanatory variables. The regression is estimated separately for men and women. The coefficients on the experience dummies are the annual growth rates of earnings for lawyers (male and female) that we use in equation (1.19).

49 Each yearly Italian LFS is a representative cross-section of the Italian population: unfortunately, it does not report the wages of self-employed individuals, including lawyers. We estimate the experience profiles in the same way they are estimated for lawyers (see footnote 48), namely via a simple OLS regression with log earnings as a dependent variable and year dummies and experience dummies (one per each year of experience) as explanatory variables. We estimate one regression for each gender and the coefficients on the experience dummies are the annual growth rates of earnings for non-lawyers that we use in equation (1.19).

Figure 1.14: Wage-experience profiles for lawyers and non-lawyers by gender



Source: Own calculations on ISTAT LFS 2009-2018 and Cassa Forense data.

Results are similar to those in our main simulations of Section 1.6. Only a few minor differences are worth noticing. Consistent with the notion that individuals make forward-looking decisions, Panel A suggests that considering lifetime earnings allows the model to fit the data a little better, especially at higher deciles of GPA. Panel B now shows a slightly stronger effect of eliminating connections in the exam than in our baseline simulations. Evidently, considering longer careers makes the returns to the legal profession larger, especially for individuals with higher GPA, who already start off with higher earnings. One potential limitation is that the experience profiles might be different for connected and unconnected individuals and we, unfortunately, have no information about this. We do not expect particularly large differences outside the legal profession and it is hard to say whether the earnings of connected lawyers would grow faster or slower than those of their unconnected colleagues.

Given how we incorporated the earnings profiles in our simulation, it is not surprising to find that they do not change much the effect of eliminating the role of connections in earnings (see Panel C). In equation (1.19), the differential growth rates simply enter linearly and are unaffected by connections. Hence, the only reason why they might influence occupational choices is when expected returns are multiplied by the probability of passing the bar exam, which does not change in the simulation shown in Panel C.

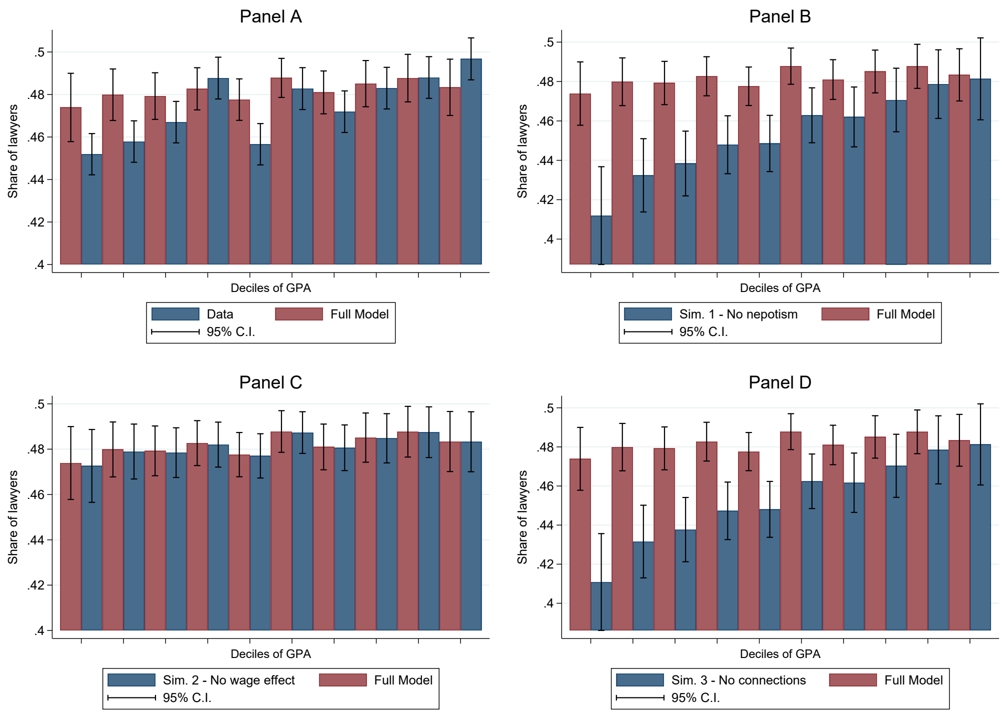
Eventually, and as in our baseline simulation, eliminating the role of connections both in earnings and in the exam yields very similar results as when eliminating them only in the exam (Panel D).

1.7.3 Sample composition

In our main analysis of Section 1.5, we estimate all the equations of the model using the same sample of observations with valid information on all the variables required for each and every equation. This approach allowed us to produce results that could be easily compared across equations, but it might also generate doubts about the implications of sample selection. In this section, we replicate our main estimates using the largest available sample for each equation separately.

Table 1.9 is the equivalent of Table 1.3 implemented on the largest available sample of individuals for whom information on our human capital indicators (and controls) is available. Despite the large difference in sample sizes, results remain very similar, both qualitatively and quantitatively. The large difference in samples arises because, when we restrict the analysis to the common sample, we need to drop several individuals for whom we have no information on earnings and apprenticeship. These variables are gathered via the post-graduation surveys,

Figure 1.15: Counterfactual simulations with wage-experience profiles



Notes: The Figure reports the results of the counterfactual simulation exercises with earnings' yearly growth rates described in Section 1.7.2. Panel A reports the baseline results of the empirical model. Panel B reports the results of the simulations with no connections at the exam stage. Panel C reports the results of the simulations with no connections at the earnings stage. Panel D reports the results of the simulations with no connections at any stage.

whereas most of the information needed to produce the estimates in Table 1.9 comes directly from administrative archives, where the issue of missing data is minimal.

Table 1.10 reproduces a similar exercise with reference to Table 1.5 in our main results of Section 1.5. The number of available observations is smaller than in the previous Table 1.9 because we now need to use information about whether the individuals have ever done a professional apprenticeship. Yet, sample size is substantially larger than in the common sample and some important differences in the estimates are present. For example, connections now appear to affect not only the probability of passing the bar exam but also the decision to undertake an apprenticeship, although the magnitude of the latter coefficient is less than half the former. Moreover, the effect of GPA on the exam is now more distinctively different between connected and unconnected individuals. Overall, our main results are confirmed.

Finally, in Table 1.11 we look at earnings, expanding the size of the sample as much as possible for each equation. Once again, results are comparable to those

Table 1.9: Occupation-specific human capital with largest possible sample

Dep. variable= GPA[a]	(1)	(2)	(3)	(4)
High school grade[b]	0.405***	0.405***	0.405***	0.405***
	(0.004)	(0.004)	(0.004)	(0.004)
1=connections[c]	-0.007	-	-	-
	(0.011)			
1= few connections[d]	-	-0.005	-	-
		(0.010)		
1= many connections[d]	-	-0.020	-	-
		(0.014)		
Number of connections	-	-	-0.001***	-0.002***
			(0.000)	(0.001)
Number of connections^2	-	-	-	-0.000
				(0.000)
1=female	0.106***	0.106***	0.106***	0.106***
	(0.008)	(0.008)	(0.008)	(0.008)
1=graduate parent[e]	0.150***	0.150***	0.150***	0.150***
	(0.008)	(0.008)	(0.008)	(0.008)
Observations	46,619	46,619	46,619	46,619

[a] Average grade in all exams at the law school. Standardised within each university. [b] Standardised over the sample. [c] 1=some connections; 0=no connections. [d] few = 1-3 ; many = 4+. [e] At least one parent with university degree.
All regressions include fixed effects for university, district, year of graduation, log size of district and log name frequency in district. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

reported in our main analysis.

1.8 Conclusions

The available evidence indicates that occupational regulation very often fails to improve the quality of professionals and the services they provide. Whereas this finding is well established, much less is known about the reasons for such a blatant failure. Of course, not knowing the reasons why occupational regulation so often fails, it is hard to offer policy advice.

In this paper, we provide what we believe to be the first systematic analysis of the mechanism by which occupational licensing selects professionals, and we highlight where and how the system breaks down. Our results suggest that the problem lies with the strong degree of intergenerational transmission of occupations that, while being a general phenomenon, is also particularly relevant in the presence of professional licensing.

Table 1.10: Probabilities of apprenticeship and exam with largest possible sample

	Probability of	
	doing an apprenticeship	passing the exam
	$P(T_i = 1 \mid Z_{ir})$	$P(E_i = 1 \mid T_i = 1, Z_{ir})$
GPA[a]	0.128***	0.004
	(0.016)	(0.013)
1=connections[b]	0.067**	0.140***
	(0.027)	(0.021)
GPA × [1=connections]	-0.034*	-0.080***
	(0.019)	(0.015)
High school grade[c]	-0.074***	-0.037***
	(0.012)	(0.009)
1=female	0.053***	-0.084***
	(0.021)	(0.016)
1=graduate parent[d]	0.035	-0.047***
	(0.023)	(0.015)
1=parent(s) in high-ranked occup.[e]	0.089***	-
	(0.022)	
Grading district FE[f]	No	Yes
Observations	38,259	34,873

[a] Average grade in all exams at the law school. Standardised within each university. [b] 1=some connections; 0=no connections. [c] Standardised over the sample. [d] At least one parent with university degree. [e] At least one parent employed as professional, entrepreneur or manager. [f] Fixed effects for the district of exam correction.
All regressions include fixed effects for university, district, year of graduation, log size of district and log name frequency in district. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Of course, our findings are specific to the context that we analyse, namely that of licensed lawyers in Italy, and they may not easily generalise to other settings. Nevertheless, the institutional environment of the legal profession in Italy is relatively standard for intellectual liberal professions in most industrialised countries. Beyond lawyers, these liberal professions include accountants, notaries, architects, and pharmacists, among others. Hence, we believe that our work can be very informative for a large and important set of regulated professions around the world. For example, Koumenta and Pagliero (2018) report that in the European Union, approximately one-quarter of the self-employed and a similar share of all graduates work in a regulated profession.

Our analysis offers insights that can be immediately useful for policy interventions. We show that system malfunctions are mostly concentrated in the entry exam, which assigns an important role to incumbent professionals and does not guarantee the complete anonymity of the candidates. Incumbent lawyers might

Table 1.11: Lawyer and non-lawyer earnings with largest possible sample

	Lawyer earnings	Non-lawyer earnings	Selection
	y_i^L	y_i^0	$P(L_i = 1 \vert Z_{ir})$
GPA[a]	0.250***	0.126***	0.054***
	(0.040)	(0.047)	(0.013)
1=connections[b]	0.192***	-0.129	0.118***
	(0.057)	(0.089)	(0.021)
GPA × [1=connections]	-0.130***	0.040	-0.087***
	(0.046)	(0.058)	(0.016)
High school grade[c]	0.058**	0.108***	-0.012
	(0.024)	(0.035)	(0.009)
1=female	-0.629***	-0.658***	-0.094***
	(0.045)	(0.056)	(0.016)
1=graduate parent[d]	0.049	0.026	-0.048***
	(0.043)	(0.054)	(0.017)
Exclusion restrictions:			
1=parent(s) in high-ranked occup.[e]	-	-	0.038***
			(0.016)
Grading district FE[f]	No	No	Yes
Observations	36,778	34,022	36,778

[a] Average grade in all exams at the law school. Standardised within each university. [b] 1=some connections; 0=no connections. [c] Standardised over the sample. [d] At least one parent with university degree. [e] At least one parent employed as professional, entrepreneur or manager. [f] Fixed effects for the district of exam correction.

All regressions include fixed effects for university, district, year of graduation, log size of district and log name frequency in district. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

have an interest in facilitating connected candidates and they might be able to do so lawfully and possibly even unconsciously. For example, by statistical discrimination, commissioners might explicitly or implicitly assume that young lawyers coming from successful dynasties of professionals are better than others. Hence, any intervention that might preserve the anonymous identity of the candidates and limit or regulate the role of incumbents could have potentially important effects on selection.

In the specific case of Italy, one could change the composition of the local commissions and avoid having lawyers of one district interviewing candidates in the same district. For example, the random allocation of districts could be extended to commissioners. A more drastic solution would be to abolish the oral exam completely.

A number of important avenues remain open for future research. Among the most important ones is the investigation of output quality. The ultimate aim of the regulation is to guarantee the quality of services that are offered on the market, whereas our analysis focuses on the quality of providers. Measuring quality is a notoriously difficult task and it is already an important achievement that we were able to measure some dimensions of input quality in this paper. Measuring output quality is even more challenging and Anderson et al. (2020) is the only paper we are aware of that addresses it. In addition, we transparently acknowledge that our measure of quality or competence is imperfect, and improving it would be a very welcome development that could lead to a better understanding of the process of professional human capital accumulation.

CHAPTER 2

Self-Selection, University Courses and Returns to Advanced Degrees

Higher education often requires choosing a bachelor's and a master's degree, yet we know little about the returns of these combined choices and the role of courses in different disciplines. This paper addresses this gap using detailed data on Italian graduates and university programs. I study the labor market returns to combinations of bachelor's and master's degrees and investigate how the characteristics of the curriculum affect outcomes. I exploit exogenous variation in access to bachelor's and master's degrees to causally estimate the returns to 43 combinations of degrees. I organize the data in a nested model with exogenous variation in admission requirements and investigate the preference profile of the sample through policy simulations that shift such requirements. I then relate the estimated returns to the academic curriculum of degrees to examine the role of quantitative education. I contribute to the literature on returns to advanced degrees by incorporating master's degrees in the discussion on how higher education affects outcomes and providing evidence on the characteristics of curricula that are positively related to labor market returns. I find that returns to degree combinations vary substantially even for combinations of degrees with the same bachelor's, suggesting that both types of programs require consideration. Combinations of degrees in different disciplines relate positively to economic outcomes, while combinations in the same field perform worse. Successful combinations have little non-quantitative education in the master's, and quantitative courses alone do not explain higher returns.[1]

[1]I thank Michele Pellizzari, Giacomo De Giorgi, and Peter Arcidiacono for their guidance and support. Aleksey Tetenov, Edwin Leuven and Arnaud Maurel provided valuable feedback,

2.1 Introduction

The literature on the returns to education is currently active on the issue of university degrees. Recent evidence suggests that alternative choices of degrees can have significant implications on labor market outcomes (Altonji and Zimmerman, 2018; Hastings et al., 2013; Kirkeboen et al., 2016; Altonji et al., 2016, 2012). A critical element in this debate that has so far gone rather unnoticed is that, within degrees, there is substantial heterogeneity in the amount of instruction across different disciplines. For example, a typical degree in economics requires a sizable number of classes in law, statistics, and math in addition to courses in economics. In this paper, I investigate the labor market value of university degrees by combining administrative data covering almost the entire universe of university graduates in Italy with purposely collected detailed information on the disciplinary content of all university programs. The data contains information on the number of compulsory classes required in each program and each class is associated with one discipline. I develop a methodology to causally estimate the labor market returns to each university program and I analyze the disciplinary content of programs with high and low returns.

I carry out the empirical exercise in the context of Italy, where most students enroll in a 2-year master's program after a 3-year bachelor's. Since the early 2000s, this is the harmonized structure of university programs across the European Union. Compared to other studies estimating the returns to degrees, this setting poses the additional empirical challenge of modeling the sequential choice of bachelor's and master's, both of which can be in several disciplines. I develop a novel methodology to estimate the returns to any combination of bachelor's and master's programs using the information on the strictness of entry requirements at both levels. In particular, for master's degrees, I have collected information about the credit requirements to enroll in any master's conditional on the previous bachelor's. For example, accessing an engineering master's from a literature bachelor's requires the acquisition of additional credits in math. I exploit this information to generate variation in the choices of bachelor's and master's that is plausibly exogenous to labor market outcomes. I organize it in a nested model in which agents first choose a bachelor's program, then, conditional on the bachelor's, choose a master's. Of course, I also allow for the choice of not doing a master's.

Several findings emerge from the analysis of 43 labor market returns to com-

along with the seminar participants at the University of Geneva, Duke University, Gerzensee Alumni Conference, SasCa PhD Conference, and Rare Voices in Economics conference. I thank Silvia Ghiselli and the AlmaLaurea research team for their expertise and hospitality throughout this project and their help accessing their resources. All mistakes are my own.

binations of bachelor's and master's degrees. First, master's choices matter for outcomes. Returns vary substantially even for combinations of bachelors' and masters' with the same choice of bachelor's. Second, combining degrees from different disciplines can improve outcomes, compared to situations where individuals specialize in the same field throughout the bachelor's and master's. All the combinations of degrees associated with the best labor market returns exhibit master's degrees in different fields than the bachelor's, while not having a master's is generally associated with worse labor market outcomes. I then investigate two features of the combinations of degrees to inform on the characteristics that relate to higher payoffs. First, I measure the amount of quantitative education in each combination of degrees and find that the relationship between labor market returns and quantitative courses is slightly U-shaped. In fact, both low- and high-earning combinations of degrees exhibit high shares of quantitative education. This finding challenges the widespread belief that degrees with more STEM (Science, Technology, Engineering, and Mathematics) education benefit students and indicates one dimension to consider when analyzing policies that incentivize enrollment in STEM. Finally, I observe that high-return combinations of degrees exhibit low shares of non-quantitative education in the master's (humanities, law, education) and relatively higher shares of non-quantitative courses in the bachelor's. This breakdown by degree level (bachelor's or master's) sheds light on the importance of the timing of courses, further corroborating the centrality of master's degrees in the analysis of returns to higher education.

My findings help us better understand how university program design affects outcomes. In particular, they contribute to the policy discussion on STEM degrees by highlighting the potential pitfalls of degrees that do not appropriately balance quantitative and non-quantitative education. Crucially, my analysis establishes the importance of advanced degrees in connection to labor market outcomes and informs on their relation to undergraduate degrees. The share of the population worldwide with a master's degree has increased steadily over the past few decades. In the U.S., the number of adults with a master's degree has more than doubled since 2000, and approximately 42% of European students and 27% of U.S. students embark on a master's degree every year (EuroStat, 2022; Hanson, 2022; US Census Bureau, 2019). Furthermore, as the U.S. higher education system allows more flexibility in the choice of classes than in Europe, the central feature of this paper – that students cover a wide range of knowledge at university – is likely to be even more relevant in the U.S. Unlike Europe, where students enroll in degrees with little flexibility, students in the U.S. can wait up to two years before declaring a major.

I contribute to the literature on returns to higher education in four directions.

Altonji et al. (2012, 2016); Oreopoulos and Petronijevic (2013), and Patnaik et al. (2020) review the literature. First, I propose an identification strategy that incorporates information about the sequential structure of the choice of degrees to causally estimate labor market returns to combinations of bachelor's and master's. Recent advancements concentrate on the limitations of using OLS and assuming selection on observables. Kirkeboen et al. (2016) exploit information on applications to higher education in Norway to account for partial rankings and estimate ex-post local heterogeneous returns to undergraduate degrees. Similarly, Hastings et al. (2013) employ a research discontinuity design that exploits threshold-crossing admissions in Chile to compute local returns that account for university reputation. Both papers use the information on private rankings of fields of study to identify the causal effect of bachelor's degrees at the margin. More recently, Bleemer and Mehta (2022) use a similar regression discontinuity approach to estimate returns to economics majors, and more selective colleges (Bleemer, 2021). Structural approaches pioneered by Arcidiacono (2004) have also been used to estimate returns to bachelor's degrees. By imposing structure on decision-making, methods relying on dynamic choice modeling can elicit ex-ante returns and incorporate introspective behaviors such as switching majors and non-pecuniary factors that can only be rationalized with error terms revealed in multiple stages. Arcidiacono et al. (2011) offer an overview of the main methods.[2] Malamud (2011, 2010) focuses on timing of specialization in higher education and its related probability of switching. He finds that early specialization in higher education is related to more costly switching. Montmarquette et al. (2002) research how students choose their majors by incorporating idiosyncratic expected earnings and heterogeneous probabilities of success and find that ex-ante expected earnings are powerful determinants of choice. Beffy et al. (2012) conversely attribute most sorting to non-pecuniary factors. I contribute to this literature by proposing an identification strategy that exploits the timing of choices and exogenous variation at different stages to retrieve labor market outcomes of combinations of degrees.

Second, I contribute to the literature on advanced degrees by incorporating them in my analysis and shedding light on the labor market enhancing features of degree combinations. Altonji and Zhong (2021) analyze the returns to detailed types of graduate programs by comparing pre- and post-graduate earnings, accounting for preferences, ability, and previous college choices. They find considerable variations in returns that are strongly related to undergraduate choices. Similarly, Arcidiacono et al. (2008) estimate returns to MBAs by taking advantage

[2]Structural approaches have also been used to identify the effect of attending selective institutions (Brewer et al., 1999) and the evolution of wage returns to education over time (Ashworth et al., 2021). d'Haultfoeuille and Maurel (2013) show that non-pecuniary factors are key ex-ante determinants of higher education attendance.

of the fact that admission into such programs requires previous work experience. Altonji (1993) estimates the returns to the highest degree obtained, including five aggregated graduate school categories, and assuming that only the highest degree matters. A few papers provide estimates of the returns to graduate degrees for specific groups of fields of study: Black et al. (2003) for individuals with economics undergraduate majors, and Bhattacharya (2005); Chen and Chevalier (2012); Ketel et al. (2016) for medical degrees. Ketel et al. (2016) is the only paper on advanced degrees not to use US data, focusing on the Netherlands. This article complements this body of work by focusing on returns for individuals who immediately enroll in a master's degree, which account for about 75% of master's graduates in Italy and 15% in the US, previously excluded from Altonji and Zhong (2021)'s analysis (AlmaLaurea, 2021b). I also exploit variation in admission eligibility to master's programs to causally estimate the returns to the complete set of bachelor's and master's combinations. The additional structure and availability of exogenous variation in incentives strengthen Altonji and Zhong (2021)'s results as they allow for rich counterfactual patterns and direct estimates of returns to degree combinations.

Third, this paper relates to the growing literature on unordered treatment effects, for which returns to university degrees are a compelling application (Bhuller and Sigstad, 2022; Heckman and Pinto, 2018; Kirkeboen et al., 2016; Mountjoy, 2022). These authors realized that when choices are unordered, the treatment effect depends on individual preferences over the choice set, even if properly accounting for self-selection. In practice, unordered settings lead to multiple contrasting margins of treatment that grow exponentially with the choice set. The large number of combinations of degrees considered in this application renders the estimation of heterogeneous margins of treatment both intractable and difficult to interpret. Bhuller and Sigstad (2022) propose an IV method to obtain economically relevant treatment effects that are averages across all heterogeneous margins.[3] This project is uniquely affected by a weak instrument problem that emerges in 2SLS estimation and that stems from the large number of endogenous regressors – the combination of undergraduate and graduate degrees – that are instrumented with the predicted probabilities of enrollment obtained with the nested model (Phillips and Gao, 2017). While the setup is close in spirit to Bhuller and Sigstad (2022), identification requires a reduced form solution to avoid using

[3]Bhuller and Sigstad (2022) propose an average monotonicity condition that requires instruments to increase the probability of treatment on average. Joint with a cross-effects condition that guarantees that instruments uniquely affect treatments, average monotonicity identifies properly estimated average treatment effects with multiple unordered treatments in 2SLS. In practice, their model exploits a modified first stage where each instrument affects the treatment separately.

the information about the correlation between the endogenous regressors and the instruments (Chernozhukov and Hansen, 2008).

Lastly, I contribute to the literature on degree characteristics. Despite the consensus that higher education is essential to labor market success beyond ability signaling, the evidence on how degrees affect outcomes lacks a systematic approach. Biasi and Ma (2022) focus on the coverage of frontier knowledge in higher education. They find that instructors play a central role in surmounting the education-innovation gap and that students with access to such knowledge earn more after graduation. Braga et al. (2016) investigate the impact of instructors in college on labor market outcomes and conversely discover a mild effect. Deming and Noray (2020) look at the skill decay of college graduates and find that earning premia decline faster for graduates in technology-intensive fields. Acemoglu et al. (2022) find that CEOs in Denmark and the US with business education are responsible for less profit sharing with employees and claim that practices and values acquired in business school are responsible. STEM degrees, characterized by quantitative and technical education, have received considerable attention. However, even within this group of degrees, there is a lack of consensus in the characteristics that are important for labor market outcomes (Xie et al., 2015). Table B.2 in appendix B.1 substantiates this claim by comparing STEM definitions in the literature. By analyzing the impact of university courses by field of study on labor market returns, I contribute with the first systematic review of labor-enhancing degree characteristics across disciplines.

The rest of the paper is organized as follows. Section 2.2 summarizes the relevant features of the Italian higher education system and discusses its similarities with the European and U.S. context. Section 2.3 describes the theoretical framework of the analysis. It presents the stages of the model and the empirical challenges in close relation to the available data. Section 2.4 describes the main data sources on Italian graduates and university programs. Section 2.5 presents the results of all the stages of the model to obtain the labor market returns to 43 degree combinations. It also presents a policy simulation that shifts admission requirements to investigate how preferences affect enrollment at the intensive margin. Section 2.6 relates the estimated returns to program characteristics such as timing, quantitativeness, and multidisciplinarity to elicit labor market enhancing characteristics. Together, these results provide the basis for the discussion on program characteristics. Section 2.7 concludes.

2.2 Institutional Background

Italy adheres to the Bologna process (1999) that ensures comparability in higher education standards across the European Higher Education Area (EHEA), which comprises 48 European and Central Asian countries. Notably, this means that degrees are organized as bachelor's (three years) and master's (two years) with comparable workloads as measured by credits, the unit of academic work. According to the European Credit Transfer and Accumulation System (ECTS), one credit corresponds to 25 hours of academic work, divided between classes and individual study. One year of higher education consists of 60 credits. Admission into a master's degree is conditional upon completing a bachelor's, and students apply for admission into programs with different fields of study. Additional objectives of the Bologna process are the automatic recognition of degrees throughout the EHEA and the promotion of international student mobility.

Throughout the paper, I will use the following terminology: a *degree* is the university program that students choose to enroll in and can refer to either a bachelor's or a master's program, a *university career* is the joint choice of a bachelor's (undergraduate) and master's (graduate) degree. A university *course* is a portion of what is studied in a degree and covers an individual subject, and its unit is one *credit*. Both degrees and courses vary as several choices of *fields of study* (disciplines) are available, and the same university course can be studied across several degrees. The *academic curriculum* refers to the prescription of courses and credits that describes a degree.

For a degree to be legally recognized, it must meet considerable requirements that govern its curriculum and are expressed in terms of course content and credit amounts. During the period of the analysis that considers graduates from 2007 to 2014, there were 47 bachelor's and 99 master's degrees.[4] Some degrees are exceptionally organized as single-cycle degrees that last five or six years and confer a master's degree without there being a corresponding bachelor degree. These include medicine, veterinary, dentistry, architecture, law, chemical and pharmaceutical technologies, and primary education.

The academic curriculum of each degree can be described along two dimensions: the number of credits to be allocated to each course and the course content. Course content is coded homogeneously across degrees and universities for a total of 370 possible disciplines (CUN, 2000). This means that all the courses offered in higher education belong to one of the codified fields. Then, the academic curriculum

[4]The Italian higher education system also includes academic diplomas, one-year master's, doctoral programs, and vocational degrees. Only academic diplomas which have equal legal standing to a bachelor's degree are considered.

of each degree further prescribes how many credits to give to each course. For example, the code MAT-5 corresponds to calculus. A course in calculus with this code can be found in 23 bachelor's degrees and 12 master's degrees, but different credits can be associated with these courses. For each degree, more than 50% of course content and number of credits is fixed. Students can freely allocate only 10% of all credits, equivalent to approximately one class per year. The remaining credits are divided between any compulsory internships and thesis periods in varying proportions. Hence, a degree is fully described by the vector of courses and credits in each discipline. Importantly, students choose degrees with a predefined curriculum rather than courses.

For statistical precision, I group bachelor's and master's degrees into ten fields of study, described in table 2.1. The grouping is consistent with the data provider's aggregation, with slight adjustments for comparability across data sources and is further discussed in section 2.4. A detailed list of which degrees belong to which group can be found in appendix B.5.2. Throughout the paper, I will focus on university careers rather than degrees, that is, a joint choice of bachelor's and master's degree. For example, a career in economics implies both a bachelor's and master's in economics, while a career in economics and law indicates a bachelor's in economics and a master's in law.

Table 2.1: Fields of study description

Code	**Abbreviation**	**Description**
1	Agr.Vet.Geo.Bio.	Agriculture and veterinarian sciences, geology and biology
2	Arch.Eng.	Architecture and Engineering
3	Chem.Pharm.	Chemistry and Pharmacy
4	Econ.Mgmt.	Economics and Management
5	Educ.Psy.	Physical education, Teaching, Psychology
6	Law	Law
7	Lit.Lang.	Literature, Languages and Humanities
8	Health	Medicine and Health-related studies
9	Pol.Soc.	Political Sciences, Sociology and Communication
10	Sci.Stat.	Math, Physics, Natural Sciences and Statistics

Students with any secondary education diploma can access university.[5] Admission into a bachelor degree can either be regulated at the national level – as it is the case with all health-related degrees, veterinary, architecture, and primary

[5] Until the late 1960s, only students with the most academic-oriented with high-school diplomas could access university. See Bianchi (2020) and Bianchi and Giorcelli (2020) for the evaluation of the reforms that expanded access to higher education to all high-school graduates.

education – or at the university level. As universities cannot significantly differentiate their programs in terms of content, when possible they use selection criteria to attract students. This characteristic will be exploited for identification, as explained in sections 2.3 and 2.4. Admission into a master's degree is conditional on having completed a bachelor's and it also typically requires the fulfillment of curricular prerequisites, conditions on the bachelor's graduation grade, and interviews. Curricular prerequisites are defined as credits in mandated courses. For example, to enroll in a master's in economics, a student must have completed 53 credits in economics, statistics, and other social sciences during the bachelor's. Tuition varies depending on the degree, the university, and family income. About one third of students do not pay any tuition because of low family income. The average annual fee for the other students is around 1,500 euros (1,628 euros in 2019. Commission/EACEA/Eurydice (2020)). Other benefits, such as housing and meal vouchers, are allocated at the regional level depending on income and merit. Private universities charge higher tuition, usually between 10 and 15 thousand euros per year for an undergraduate program, and they govern their own merit- and need-based grants. All higher education regulations in terms of degree types, academic curricula, and admission apply to both private and public institutions. In years 2011 and 2012, only 8.17% of all university students were enrolled in private institutions (ISTAT, 2021).

2.3 Theoretical Framework

The empirical exercise in this paper consists of two stages. First, I estimate labor market returns to university careers. This is done through a nested random utility model that accounts for timing of choices and self-selection. In fact, not accounting for the choice structure leads to biased results as students self-select into careers based on observed and unobserved characteristics, and choices are unordered. Then, I use the information about the disciplinary content of degrees to investigate various policy-relevant questions on degree design. I ask whether the academic careers with the highest labor market returns are also the ones with the most quantitative or STEM content. Moreover, I check whether specializing early (during the bachelor's) or late (during the master's) in a given discipline is associated with high labor market returns. Finally, I also study whether multidisciplinarity, i.e. doing a master's in a different discipline from one's bachelor's, pays off in terms of outcomes.

This section focuses on the first part of the empirical exercise and illustrates how I retrieve the labor market returns to university careers. Section 2.3.1 lays out the methods used to obtain the probabilities of enrollment into any university career that exploit the timing of choices and exclusion restrictions. Section 2.3.2 illustrates how the probabilities of enrollment engage with a simple function of labor market outcomes (employment and wages) to obtain causal returns to university careers. The theoretical framework is set up in close relation with the available data, discussed in section 2.4.

2.3.1 Sequential Choices of Bachelor's and Master's

Here, I discuss the estimation procedure that leverages a nested logit model and exclusion restrictions to identify the individual probability of enrolling in any university career. The modeling choice stems from its choice-theoretic connection to dynamic discrete choice problems, where the intuition of these methods is that conditional on observed state variables, one can express future utility terms as functions of the probabilities that such choices occur (Hotz and Miller, 1993). Sequential choice problems with discrete unordered choices can be estimated with conditional choice probability (CCP) estimators that are brought to the data with nested logit models under the assumption of generalized extreme valued (GEV) distributed errors (Arcidiacono et al., 2011). The model allows for unobserved determinants of the choices to be correlated across nests (Hoffman and Duncan, 1988; McFadden, 1974; Montmarquette et al., 2002; Bamberger, 1987) and is implemented sequentially for tractability (McFadden, 1984; Amemiya, 1985).

One important feature of my analysis - contrary for example to Montmarquette et al. (2002) - is that I do not model the alternative outcome of not choosing a bachelor's degree. Thus, the underlying assumption is that a student who is not admitted to their preferred degree will opt for another one, rather than not studying at university. This assumption is mostly dictated by the nature of my data but it is reasonable in a public, geographically widespread, and inexpensive higher education system such as the Italian one.

Let $i \in I$ denote individuals, $j \in B$ denote a choice of bachelor's degree with $\dim(B) = L \in \mathbb{N}$, $m \in M$ denote a choice of master's degree or no master with $\dim(M) = L + 1$, such that $jm \in B \times M$ denotes a university career and $\dim(B \times M) = L(L + 1)$. The timing is as follows: in the first period, the individual must choose a bachelor's degree; in the second period, they must choose a master's degree conditional on their choice of bachelor's; ultimately, the student enters the labor market where outcomes will depend on her choice of education. In the second period, students may additionally choose not to enroll in a master's degree, thus entering the labor market directly.

In the first period, a student $i \in I$ chooses a bachelor $j \in B$. The choice will depend on characteristics that vary with the student, as well as characteristics that vary with the choice. The probability that a student i chooses a bachelor j is given by

$$P_{ij} = \frac{\exp\{X_i\beta_j + Z_{ij}\lambda_j\}}{\sum_{k=1}^{B}\exp\{X_i\beta_k + Z_{ik}\lambda_k\}} \tag{2.1}$$

where X_i is a matrix of characteristics that vary with the individual (gender, family background, general ability) and Z_{ij} is a matrix of characteristics that vary both with the individual and the choice of bachelor's (a composite measure of selectivity of admission requirements and distance to college for all bachelors'). The variation in Z_{ij} ensures that the vector of probabilities for every counterfactual bachelor degree and individual $P_{ij}\ \forall\ j \in B$ can be computed.[6]

The second nest of the model captures the choice of master's degree $m \in M$ conditional on a previous choice of bachelor's j, where M also includes the choice of not enrolling in a master's and entering the labor market directly. Similar to the choice of bachelor's, the probability that a student i chooses master m conditional on bachelor j is given by

$$[P_{im} \mid j] \;=\; \frac{\exp\{X_i\beta_m + Z_{im}\lambda_m\}}{\sum_{n=1}^{M}\exp\{X_i\beta_n + Z_{in}\lambda_n\}} \tag{2.2}$$

where X_i is defined as before and Z_{im} is a matrix of characteristics that vary both with the individual and the choice of master (factors that determine the individual's eligibility for enrollment into each master's degree), conditional on the previous choice of bachelor's j. In practice, I observe enrollment constraints for each master's that vary with the previous choice of bachelor's and can be reconstructed for each jm pair. Once again, the variation in Z_{im} ensures that the probability of choosing every counterfactual master's degree can be computed, $P_{im} \mid j\ \forall\ j \in B,\ m \in M$.
Then, the probability of enrolling in any career accounting for self-selection follows from equations (2.1) and (2.2) is given by

$$P_{ijm} = P_{ij} \times [P_{im} \mid j] \quad \forall\ j \in B,\ m \in M \tag{2.3}$$

where

$$\sum_{j=1}^{B}\sum_{m=1}^{M} P_{ijm} = 1 \quad \forall\ i.$$

[6]For clarity, I omit additional covariates throughout this section such as cohort and geography fixed effects and other controls. Section 2.5 addresses them in detail.

P_{ijm} is the predicted probability of enrollment into degree combination jm that credibly accounts for self-selection since equations (2.1) and (2.2) account for general ability and family background inter alia, as well as exogenous variation in the ease of access into degrees. Importantly, the variation in matrices Z_{ij} and Z_{im} allows for the computation of the probability of choosing every counterfactual degree-pair, overcoming the main problem in the computation of returns to degrees, which is the lack of sufficient instrumental variables to account for all possible (endogenous) choices. In principle, any number of returns to degree-pairs can be computed with this approach, as long as there is sufficient variation in Z_{ij} and Z_{im}. In practice, the estimation of the nonlinear equations (2.1) and (2.2) with maximum likelihood and the relatively high dimensionality of X_i, Z_{ij} and Z_{im} imposes constraints on the number of probabilities P_{ijm} that can be estimated. This means that university careers which are infrequently chosen may be difficult to estimate.

2.3.2 Returns to University Careers

I exploit probabilities P_{ijm} to identify the effect of career $(j,\, m)$ on labor market outcomes in a simple function

$$y_i = X_i\beta + \sum_{j=1}^{B}\sum_{m=1}^{M} P_{ijm}\alpha_{jm} + \epsilon_i \tag{2.4}$$

where y_i is the labor market outcome of interest (log wages, employment), X_i is a vector of individual characteristics and controls (gender, family background, high school grades), and α_{jm} denotes the effect of the potential treatment (careers) on outcomes. I interpret α_{jm} as the effect of university career jm on the labor market outcome y_i. These coefficients represent my object of interest as they will then be used to investigate the relationship between degree characteristics and economic outcomes in section 2.6. The empirical specification will additionally include rich sets of fixed effects (cohort, geography), detailed in section 2.5. I resort to this functional form to address three challenges to identification: self-selection on unobserved characteristics, the unordered nature of university careers, and the considerable number of choices.

To best understand the implications of these three challenges, I compare equation (2.4) with the extreme case of no-self selection into university careers on unobserved characteristics. In this case, the simple OLS regression

$$y_i = X_i\beta + \sum_{j=1}^{B}\sum_{m=1}^{M} D_{ijm}\gamma_{jm} + u_i \tag{2.5}$$

would return the effect γ_{jm} of career (treatment) D_{ijm} on outcome y_i relative to some excluded category D_{i0}, conditional on observed individual characteristics X_i, and γ_{jm} and α_{jm} would coincide. Clearly, any attempt to estimate equation (2.5) directly will result in strongly biased results as we expect students to enroll in careers based on unobserved characteristics. I address self-selection in equation (2.4) by leveraging exclusion restrictions Z_{ij} and Z_{im} in equations (2.1)-(2.3) to compute P_{ijm}.[7]

The second – more nuanced – challenge stems from the unordered nature of university careers. This equally affects equations (2.4) and (2.5) as it concerns the identification of counterfactuals, that is, the benchmark (omitted) choice against which I measure the effect of each career. Importantly, when choices are unordered, the omitted category is non-neutral and should represent at least the second preferred option or lack of treatment (Kirkeboen et al., 2016; Bhuller and Sigstad, 2022; Heckman and Pinto, 2018). To illustrate this point, consider a simplified setting with only three choices – math (M), humanities (H), and economics (E) – and two observationally identical students who enroll in economics. In this case, the effect of studying economics may not be identifiable without further information on partial rankings if absent the choice of economics, the two students choose to enroll in different degrees. To address this issue, I assume that the excluded category D_{i0} (and consequently P_{i0}) is a good proxy of lack of treatment. Section 2.5 describes the omitted category and its implications. In the example, the effect of studying economics may be heterogeneous or even contrasting depending on the choices the students would make if their preferred option were not available. A student who alternatively chooses humanities might benefit from studying economics if $y_E > y_H$, ceteris paribus, while a student who alternatively chooses math might suffer if $y_E < y_M$.[8] This is the case in all unordered settings, with the number of heterogeneous margins of treatment increasing with the number of options. Given the high number of combinations of bachelor's and master's degrees, this setting allows for up to $L^4 + 2L^3 - L$ margins of treatment, which are unlikely to be of economic relevance.[9] The aggregation of the numerous heterogeneous margins to obtain meaningful effects requires proper weighting, which relies on two conditions: that the instruments affect choices monotonically on average,

[7]As equations (2.1)-(2.4) are estimated sequentially, I obtain the standard errors of α_{jm} through pairwise bootstrapping, further discussed in section 2.5.

[8]See Mountjoy (2022) for a thorough discussion on contrasting margins of treatment.

[9]$\dim(B \times M) = L(L+1)$. Then, the number of possible margins of treatment is equal to $L(L+1) \cdot (L(L+1)-1) = L^4+2L^3-L$. In comparison, Mountjoy (2022) focuses on three possible treatments and six contrasting margins. Similarly, a practical application of Heckman and Pinto (2018) who also focuses on unordered treatments identifies a subset of interesting margins (Braccioli et al., 2022). Heckman et al. (2006); Heckman and Urzua (2010) also investigate the constraints imposed by settings with unordered treatments.

and that they do not cross-contaminate choices (Bhuller and Sigstad, 2022). Lack of cross-contamination implies that given a university career jm', any instrument $P_{jm'}$ is uniquely relevant for treatment $D_{jm'}$. This means that if instrument $P_{jm'}$ does not induce agent i into treatment jm', it cannot impact treatment $jm'' \neq jm'$ in any way that changes behavior.[10] The stepwise estimation of P_{ijm} with equations (2.1)-(2.3) allows for rich substitution patterns within which it is reasonable to assume average monotonicity of Z_{ij} and Z_{im} with respect to choices j and m (i.e. marginally shifting the admission requirements to one degree j affects choices monotonically on average within each career jm). The nested setup also reduces the chances of cross-contamination between P_{ijm} and D_{ijm} as variation in admission requirements is allowed to simultaneously affect many outcomes. Taken together, these conditions are necessary to ensure that instruments induce changes in treatment uptake in a single, threshold-crossing manner even in an unordered setting (Vytlacil, 2002; Heckman and Pinto, 2018).

The third challenge addressed by equation (2.4) pertains to the number of career effects α_{jm} of interest which can be as high as $L(L+1)$. By exploiting the reduced form, I do not need to leverage the correlation between P_{ijm} and D_{ijm} for identification, as would be the case in a two-staged least squares setting where P_{ijm} serves as an instrument for treatment D_{ijm} (Chernozhukov and Hansen, 2008). To understand why the dimension of α_{jm} can be an issue, consider the following modified 2SLS with a simplified first stage regression proposed by Bhuller and Sigstad (2022) to ensure the proper weighting of heterogeneous margins

$$D_{ijm'} = X_i\beta_{jm'} + P_{ijm'}\varphi_{jm'} + v_{ijm'} \tag{2.6}$$

for any arbitrary treatment $jm' \in B \times M$, such that treatment effects ψ_{jm} are calculated as

$$y_i = X_i\beta + \sum_{j=1}^{B}\sum_{m=1}^{M} \hat{D}_{ijm}\psi_{jm} + u_i.$$

Equation 2.6 differs from the first-stage equation in a standard 2SLS framework because only the instrument pertaining to the treatment on the left-hand side is included, i.e., $\varphi_{jm'}$ is a scalar.[11] As the number of endogenous choices

[10] Lack of proper weighting due to cross-contamination of instruments may lead to severe misrepresentation of the treatment effects. In extreme cases, cross-contamination of instruments may result in a negative average treatment effect of career jm even if all heterogeneous margins of treatment are positive (Bhuller and Sigstad, 2022).

[11] Standard 2SLS requires the estimation of $B \times M$ first stage equations for every career (j, m):

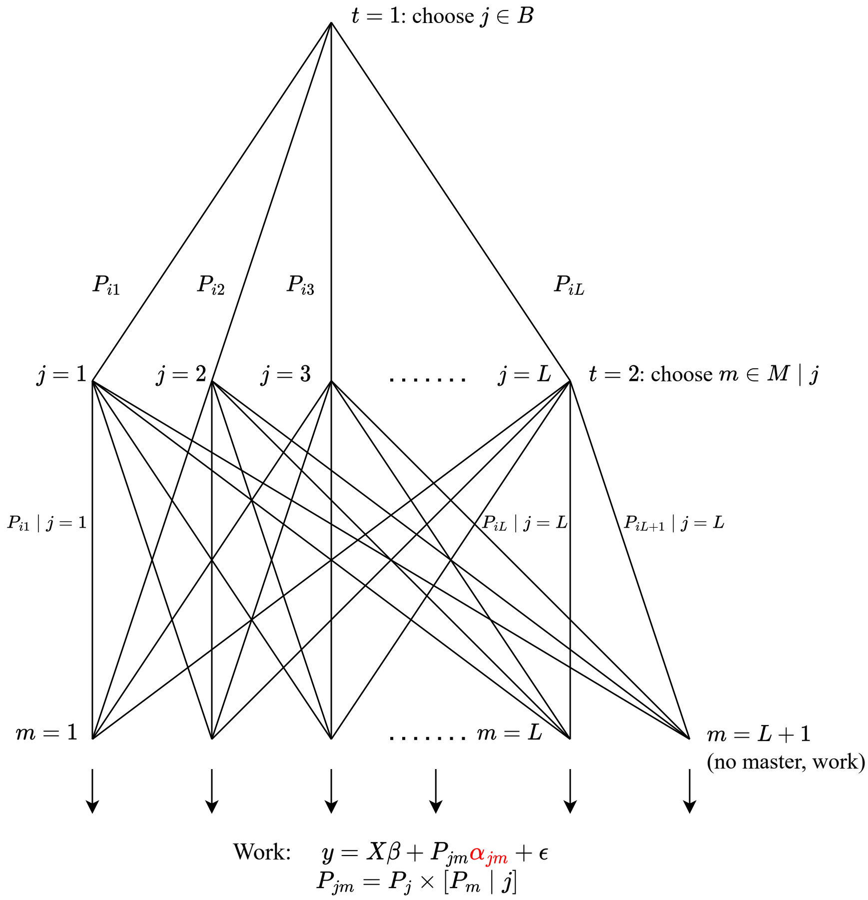
$$D_{ijm} = X_i\beta_{jm} + \sum_{k=1}^{B}\sum_{n=1}^{M} P_{ikn}\varphi_{kn} + \nu_{ijm}.$$

increases, it becomes increasingly plausible that at least some instrument P_{ijm} is not sufficiently correlated with treatment D_{ijm} even when it is relevant, thus incurring a weak instrument problem. When probabilities P_{ijm} are jointly strongly relevant, the reduced form coefficients α_{jm} asymptotically identify treatment effects ψ_{jm} (Chernozhukov and Hansen, 2008; Phillips and Gao, 2017; Crudu et al., 2021; Mikusheva and Sun, 2022). I discuss the implications of this assumption in section 2.5.1.1.

By addressing these three empirical challenges, I can interpret α_{jm} as the average treatment effect of enrolling in career jm. One alternative interpretation of α_{jm} that does not require the IV-equivalence assumptions on single threshold-crossing to hold relies on the structural interpretation of the nested model in section 2.3.1 as a dynamic discrete choice model (Arcidiacono et al., 2011). In this case, α_{jm} is the future utility term of a particular choice or the ex-ante treatment effect. The assumptions that support this interpretation require us to believe equations (2.1) and (2.2) accurately incorporate the determinants of the decision-making process of university career. Indeed, a wealth of sophisticated structural models has exploited this type of information to understand how students make schooling decisions (Arcidiacono, 2004; Ashworth et al., 2021; d'Haultfoeuille and Maurel, 2013). Lastly, α_{jm} can always be interpreted as the labor market effect of shifts in the potential treatment driven by changes in the admission requirements Z_{ij} and Z_{im}. In this setting, all instruments are jointly strongly relevant, increases in instruments P_{ijm} increase the probability of treatment D_{ijm} for all careers jm, and the nested model suggests that P_{ijm} should only affect D_{ijm}. For these reasons, I interpret α_{jm} as equivalent to IV estimates.

Figure 2.1 summarizes the timing and structure of the choice of university careers and how it integrates with the estimation of labor market outcomes α_{jm}. Exclusion restrictions that mimic admission procedures at each stage allow for the computation of counterfactual probabilities of choosing any alternative university career, partialling out the self-selection due to preferences, ability, and family background. Such counterfactual probabilities are then used as instruments for university career treatments to retrieve the causal effect of the choice of bachelor's and master's on labor market outcomes. The exploitation of timing to retrieve valid instruments allows for rich substitution patterns. An additional advantage of modeling the decision-making process explicitly is that, unlike standard 2SLS settings, it allows for students to be both forward-looking and introspective in their choices. In fact, by allowing for correlation between nests, the error term ϵ_i is allowed to be realized in multiple stages. Even though the equations of the model could be jointly estimated, the lack of certain degree combinations warrants that they be estimated sequentially. This implies that all standard errors must be

Figure 2.1: Model representation



bootstrapped to account for the method's sequential structure.

Finally, it is worth underscoring why standard 2SLS does not produce appropriate treatment effects. Not only does it allow for instruments to cross-contaminate treatments, it also imposes the estimation of a large number of irrelevant parameters which introduce significant strain on the estimator. Including irrelevant instruments on the right-hand side of the first-stage regression will decrease the precision of the estimate of the treatment effect in the second stage because it will lead to possible collinearity between instruments and inflate the standard errors of the first-stage predictions. This is especially true if – as it is the case – certain probabilities $P_{ijm'}$ are close to zero for individuals who are observed to choose $jm'' \neq jm'$.[12]

[12] Let us consider a simplified framework for presentation purposes where there are only two possible choices in each set $B = \{H, S\}$ and $M = \{H, S\}$, with H denoting "humanities" and S "science". Then $jm \in B \times M = \{HH, HS, SH, SS\}$ and the "standard" first-stage regressions

2.4 Data Sources and Summary Statistics

For my empirical analysis, I combine three data sources. The first is an administrative student-level database covering the universe of all graduates from both bachelor's and master's programs at most Italian universities, both public and private. A consortium of universities maintains this administrative archive by combining and harmonizing the original student records shared by each university. The same consortium administers surveys to all the graduates in their archives at the time of graduation and one, three, and five years later. This is my second source of data and it is individually (but anonymously) linked to the administrative records.[13] The third data source is a novel archive of administrative information about the detailed content of all university programs in Italy, including admission requirements for all bachelor's and master's programs.

2.4.1 University Graduates

My working sample considers all the individuals who graduated from 2007 to 2014, such that I observe the most recent outcomes in 2019. Eventually, I have information on 655 847 students. According to a comparison with the National Statistical Institute's (ISTAT) records, the raw sample covers between 62% and 76% of all graduates in the years of interest.[14] Several analyses carried out by the consortium suggest that the composition of their sample accurately reflects the national population of graduates over time (AlmaLaurea, 2020, 2021a). The survey data is collected online and through phone interviews. Response rates are extremely high (91%) for the first survey, administered before graduation, but remain high also for the later ones (88% across cohorts one year after graduation,

of a 2SLS model become

$$\begin{aligned}
D_{iHH} &= X_i\varphi_X^{HH} + P_{iHH}\varphi_{HH}^{HH} + P_{iHS}\varphi_{HS}^{HH} + P_{iSH}\varphi_{SH}^{HH} + P_{iSS}\varphi_{SS}^{HH} + u_{iHH} \\
D_{iHS} &= X_i\varphi_X^{HS} + P_{iHH}\varphi_{HH}^{HS} + P_{iHS}\varphi_{HS}^{HS} + P_{iSH}\varphi_{SH}^{HS} + P_{iSS}\varphi_{SS}^{HS} + u_{iHS} \\
D_{iSH} &= X_i\varphi_X^{SH} + P_{iHH}\varphi_{HH}^{SH} + P_{iHS}\varphi_{HS}^{SH} + P_{iSH}\varphi_{SH}^{SH} + P_{iSS}\varphi_{SS}^{SH} + u_{iSH} \\
D_{iSS} &= X_i\varphi_X^{SS} + P_{iHH}\varphi_{HH}^{SS} + P_{iHS}\varphi_{HS}^{SS} + P_{iSH}\varphi_{SH}^{SS} + P_{iSS}\varphi_{SS}^{SS} + u_{iSS}.
\end{aligned}$$

As this approach forces the estimation of $(B \times M - 1)^2$ irrelevant parameters, there is a serious concern of overidentification in the first stage, which is exacerbated if some P_{ijm} is small and aggravates any weak instrument bias.

[13]The AlmaLaurea Inter-University Consortium collaborates with Italian universities and the Ministry of University and Research (MUR) to monitor the labor market outcomes of Italian graduates and help match graduates with employers. Universities adhere to the consortium in different years, with 80 out of 96 universities participating in 2022. The full list of participating universities can be found in appendix B.5. Access to their resources is restricted.

[14]In 2007, only 46 universities of all 96 adhered to the consortium, while 64 were participating by 2014. I do not consider earlier cohorts since they only include students who graduate in July of each year, university participation was lower, and a different university system was still fading away.

81% after three years, and 75% after five years). The surveys provide information about socio-economic characteristics and labor market outcomes.

Two limitations are intrinsic to the setup. First, I only observe students who complete at least a bachelor's degree. Hence, any conclusion from the empirical analysis should be interpreted at the intensive margin. Second, I do not observe university dropouts. This is relevant for master's graduates, as it is impossible to distinguish between outmigration of bachelor's graduates to institutions outside of the consortium, and master's students who drop out. To avoid confusing the two, among bachelor's graduates without a master's degree, I only keep those who report no intention of enrolling in a master's program.[15] Second, ancillary information on local labor market conditions is not available for international students who are dropped from the main analysis. They account for less than 2% of the dataset, as most international mobility occurs through Erasmus and similar short-term exchange programs.[16]

Table 2.2 presents the distribution of individuals across university careers. Groups with fewer than 100 observations (in red) are dropped to ensure sufficient power during estimation for a total of 1 325 observations. 56 groups out of 110 contain sufficient records. 60.8% of graduates complete both a bachelor's and a master's degree. 24 433 (6.1%) of master's graduates switch disciplines after the bachelor's. This value is very conservative as it depends on the grouping of degrees in broad fields. Less conservative groupings observe switching in up to 15% of cases. Section 2.4.2 elaborates on the grouping rule.

Table 2.3 presents the descriptive statistics of the main individual characteristics, summarized by bachelor's degree. The characteristics that vary the most across fields are gender and high school type. Even though there are 62% of women in the sample, female students are under-represented in architecture and engineering (34%) and science and statistics (35%), and are over-represented in education and psychology (83%) and humanities (78%). High school types are grouped into three main categories: sciences, humanities, and other high schools, including languages, social sciences, and vocational schools. Although no high school type precludes enrollment into any degree, we remark more students with a humanities high school in literature and languages (23%) and law (34%). Students from science high schools are over-represented in life sciences, engineering,

[15] The survey asks bachelor's graduates whether they intend to enroll in a master's degree abroad, enroll in a different type of program (e.g. one-year master's), or not enroll. In addition to master's graduates, I only keep bachelor's graduates who do not intend to further enroll in higher education. Fortunately, attrition due to outmigration seems low, as only 1.4% state an intention to enroll in a master's that is not observed by the consortium.

[16] The employment rate for individuals 25-34 years old in the province of birth before enrollment into university summarizes local labor market conditions. The information is obtained from the National Statistical Institute (ISTAT).

Table 2.2: Frequency of graduates in all university careers

	Master's											
Bachelor's	No Master	AVGB	Arc.Eng.	Chem.Ph.	Ec.Mg.	Ed.Psy.	Law	Lit.Lan.	Health	Pol.Soc.	Sci.Stat.	Total
AVGB	8,387	26,316	219	59	19	180	*	41	622	29	932	36,656
Arc.Eng.	22,285	87	79,827	776	84	18	*	287	10	91	251	103,426
Chem.Ph.	3,902	118	11	20,643	*	*	*	*	260	*	18	24,923
Ec.Mg.	27,806	23	16	*	46244	123	208	67	31	1,153	459	75,993
Ed.Psy.	28,530	26	*	*	16	46,085	18	250	125	537	11	75,527
Law	8,054	*	27	*	1,466	127	46,766	84	24	1,101	13	57,514
Lit.Lan.	38,343	76	122	27	693	595	55	44,974	27	5,788	166	90,681
Health	75,743	403	29	*	16	313	*	11	28,056	50	*	104,515
Pol.Soc.	35,003	*	65	*	1,562	599	1,342	1,949	24	25,324	112	65,891
Sci.Stat.	8,597	1,014	115	183	123	15	*	60	*	160	10,529	20,721
Total	256,650	27,851	80,283	21,602	50,088	48,022	48,316	47,460	29,063	34,063	12,449	655,847

Frequencies in red denote careers that are observed for less than 100 individuals. Asterisks indicate groups with fewer than 10 individuals. All groups except 3 are chosen at least once. Total amounts do not include the less frequent choices in red. AVGB – Life Sciences, Arc.Eng. – Architecture and Engineering, Chem.Ph. – Chemistry and Pharmacy, Ec.Mg. – Economics and Management, Ed.Psy. – Education and Psychology, Lit.Lan. – Humanities, Literature and Languages, Law – Law, Health – Medicine and Health, Pol.Soc. Political and Social Sciences, Sci.Stat. – Math, Physics and Statistics.

chemistry and hard sciences. I include two measures of family background: parent education, measured as at least one parent with a college degree, and parent occupation, that is, at least one parent in a high-ranked profession, such as executive, entrepreneur, professional, or academic. Neither of these measures varies dramatically across fields. One exception is law degrees, where relatively more individuals have parents with college degrees (36%) and in high-ranked occupations (30%). I standardize high school final grades by province to account for differences in grading standards across school districts. Relatively more students with above-average high school grades enroll in engineering (62%) and hard sciences (58%). Below-average high school grades are observed in education (37%), social sciences (41%) and healthcare (42%).

The main empirical analysis focuses on two labor market outcomes: log wages and employment five years after graduation.[17] Figure 2.2 presents average wages in levels reported to 2015 Euros for the sample of the employed, which tallies 508 242 records (77%), for each academic career. Figure 2.3 shows similar summary statistics for average employment levels over the whole sample of 655 847 graduates. Both figures 2.2 and 2.3 display differences in labor market outcomes by undergraduate choice of major by comparing the solid and dashed red lines.

[17]When the outcomes are not available five years after graduation, they are imputed using the one- and three-year survey waves. The main empirical analysis includes survey-wave fixed effects to account for these differences.

Table 2.3: Description of the main individual characteristics by bachelor's field of study.

Variables	All (1)	AVGB (2)	Arch.Eng. (3)	Chem.Ph. (4)	Econ.Mg. (5)	Educ.Psy. (6)	Law (7)	Lit.Lan. (8)	Med. (9)	Pol.Soc. (10)	Sci.Stat. (11)
High School: grade (st.)	0.00	0.04	0.30	0.11	0.04	-0.30	0.08	0.10	-0.19	-0.20	0.22
	(1.000)	(0.969)	(0.954)	(0.955)	(0.991)	(0.954)	(0.981)	(0.984)	(1.021)	(0.979)	(0.997)
High School: humanities	0.15	0.13	0.08	0.18	0.07	0.13	0.34	0.23	0.14	0.16	0.06
	(0.359)	(0.337)	(0.271)	(0.380)	(0.258)	(0.342)	(0.474)	(0.423)	(0.344)	(0.365)	(0.241)
High School: science	0.39	0.52	0.55	0.57	0.36	0.27	0.32	0.26	0.42	0.29	0.52
	(0.487)	(0.499)	(0.498)	(0.495)	(0.481)	(0.445)	(0.468)	(0.441)	(0.494)	(0.452)	(0.500)
Gender (1=female)	0.62	0.60	0.34	0.69	0.54	0.83	0.63	0.78	0.68	0.69	0.36
	(0.485)	(0.490)	(0.474)	(0.463)	(0.499)	(0.378)	(0.483)	(0.414)	(0.468)	(0.464)	(0.479)
Parents: graduate	0.26	0.27	0.31	0.33	0.22	0.18	0.36	0.26	0.23	0.22	0.27
	(0.438)	(0.444)	(0.463)	(0.472)	(0.416)	(0.384)	(0.481)	(0.439)	(0.420)	(0.415)	(0.442)
Parents: high-rank occ.	0.21	0.21	0.25	0.26	0.22	0.16	0.30	0.21	0.18	0.20	0.18
	(0.410)	(0.405)	(0.431)	(0.440)	(0.411)	(0.368)	(0.459)	(0.404)	(0.388)	(0.400)	(0.383)
Observations	655847	36656	103426	24923	75993	75527	57514	90681	104515	65891	20721

Column labels: AVGB – Life Sciences, Arch.Eng. – Architecture and Engineering, Chem.Ph. – Chemistry and Pharmacy, Econ.Mg. – Economics and Management, Educ.Psy. – Education and Psychology, Lit.Lan. – Humanities, Literature and Languages, Law – Law, Med. – Medicine and Health, Pol.Soc – Political and Social Sciences, Sci.Stat. – Math, Physics and Statistics.

The figures also point to large differences in outcomes by combinations of undergraduate and graduate majors, visible by comparing the vertical bars within each subgraph. Overall, individuals without a master's degree experience worse labor market outcomes on average (first column of each subgraph). Even though these figures present unconditional means, they suggest that outcomes vary substantially across masters' choices also conditional on bachelors'.

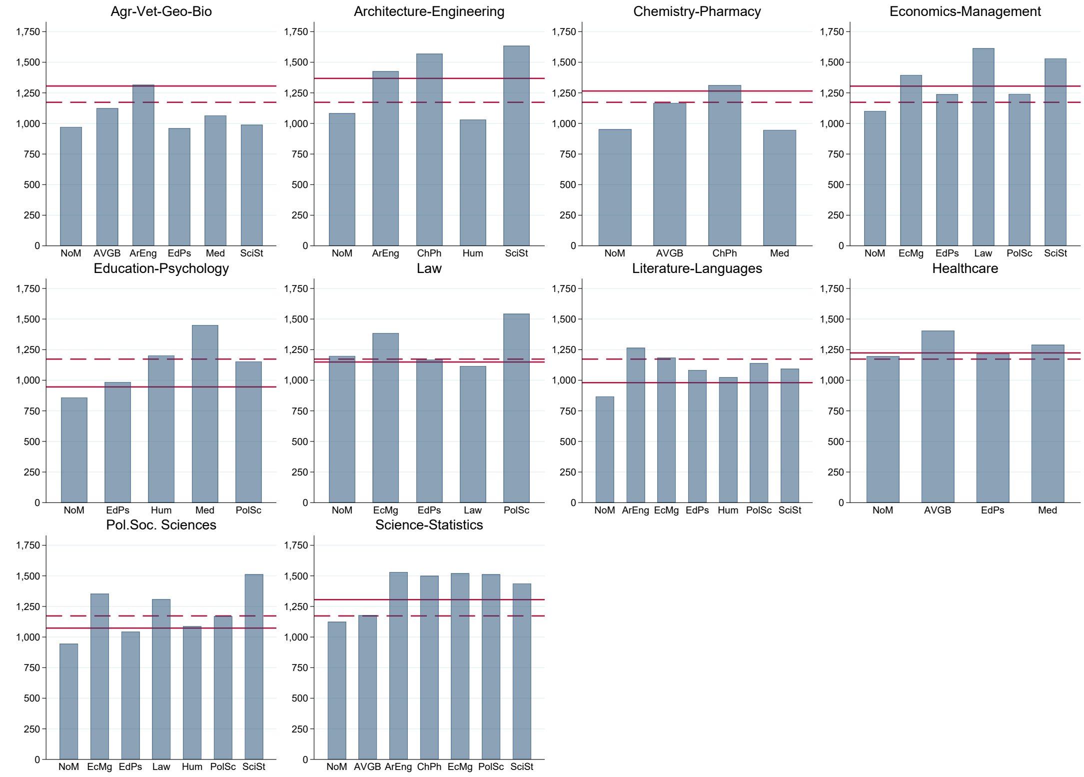
2.4.2 University Programs

I complement the student records with detailed information about the content and structure of all academic programs. The data on the content of programs combines various legal sources to reconstruct the compulsory features of degrees. The data on the structure of programs focuses on admission practices and results from a survey of all programs offered in Italy.

The data on the content of academic programs comes from two sources: the content requirements in terms of credits and courses of all 47 legally recognized bachelor's programs and 99 master's programs, and the official codes and descriptions of 370 available disciplines.[18] Crucially, I observe the disciplinary content of

[18]Law 270/2004 provides detailed information on the legal requirements that degrees must

Figure 2.2: Description of wages in 2015 Euros by academic career.



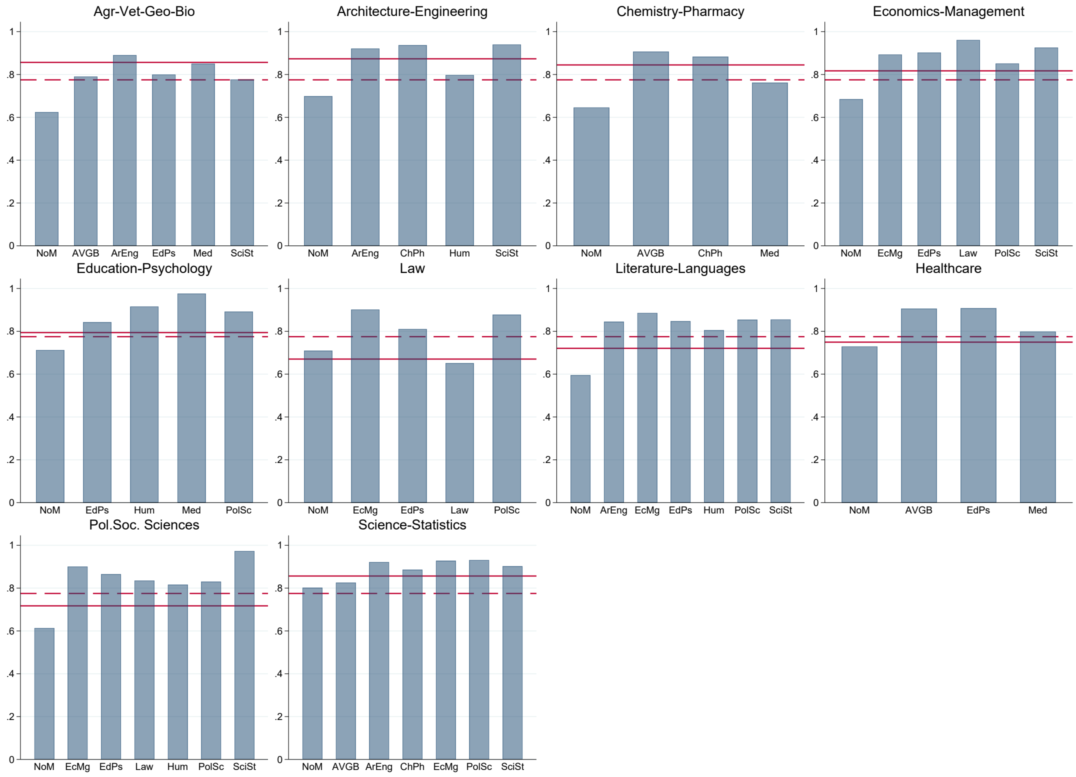
Sub-graph titles indicate the bachelor's choice, while the fields of study on the horizontal axis refer to master's choices. The solid red line represents the average wage in 2015 Euros for the subsample of individuals who share the same bachelor's choice. The dotted red line indicates the sample average. NoM – No Master, AVGB – Life Sciences, ArEng – Architecture and Engineering, ChPh – Chemistry and Pharmacy, EcMg – Economics and Management, EdPs – Education and Psychology, Hum – Humanities, Literature and Languages, Law – Law, Med – Medicine and Health, PolSc Political and Social Sciences, SciSt – Math, Physics and Statistics.

any university course independently of the institution or the degree in which it is taught. Furthermore, for each course I observe the number of credits that must be obtained to meet the program's legal requirements. I use this information to account for different levels of specialization across degrees. For example, a course in applied economics is present in 17 bachelor's programs and 33 master's programs. However, the number of required credits varies greatly, from 4 credits in a master's program in architecture to 32 credits in a bachelor's in economics.

Figure 2.4 presents a complete description of the content of bachelor's (left) and master's (right) degrees at the relevant level of aggregation, by plotting them against their academic curriculum, with the total percentage of required credits in the degree's main field of study on the diagonal. Each line represents a degree

meet. Addenda to the law have been exceptionally published over the years and are considered when relevant. The list of scientific disciplines (*settori scientifico-disciplinari*) is maintained by the Italian National University Council (CUN). The total number of disciplines has increased since the years under consideration to 384.

Figure 2.3: Description of employment by academic career.



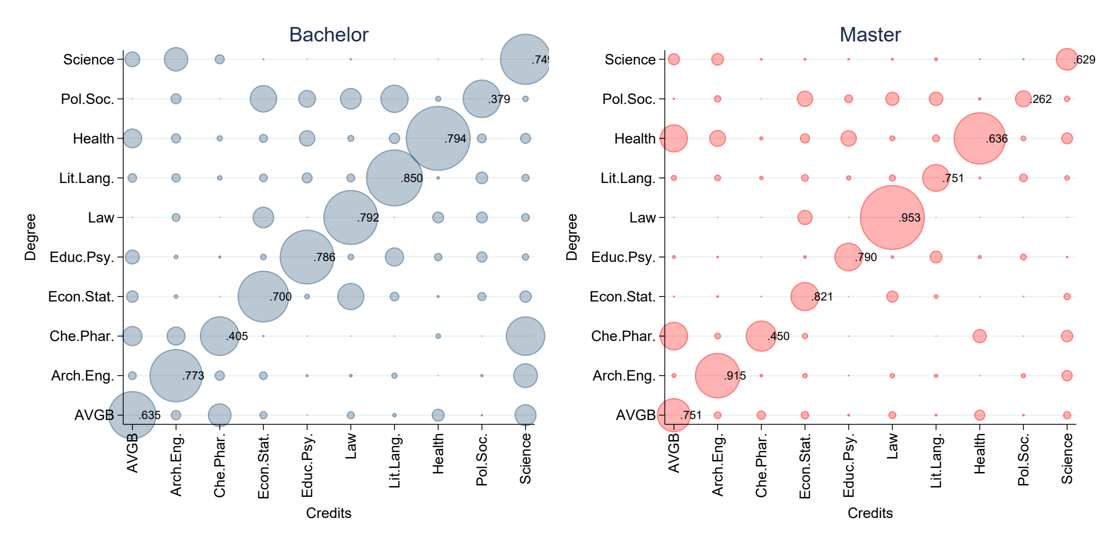
Sub-graph titles indicate the bachelor's choice, while the fields of study on the horizontal axis refer to master's choices. The solid red line represents the average level of employment for the subsample of individuals who share the same bachelor's choice. The dotted red line indicates the sample average. NoM – No Master, AVGB – Life Sciences, ArEng – Architecture and Engineering, ChPh – Chemistry and Pharmacy, EcMg – Economics and Management, EdPs – Education and Psychology, Hum – Humanities, Literature and Languages, Law – Law, Med – Medicine and Health, PolSc Political and Social Sciences, SciSt – Math, Physics and Statistics.

by averaging the content of each program that belongs to the degree grouping.[19] Indeed, there is significant off-diagonal variation, with two degrees – chemistry and pharmacy, and political and social sciences – requiring less than 50% of time studying the main discipline both at the undergraduate and graduate level. While degrees specialize slightly during the master's, with more credits in the main domain, there still is substantial education in off-diagonal fields. The grouping of degrees, described in table 2.1, serves two objectives: yield statistical precision and economically interesting results. I primarily base the grouping on that of the data provider and the Italian ministry of higher education. Infrequently chosen groups are further grouped according to the literature (table B.2 overviews some of the papers that were used) to maintain proximity in content. To further validate this approach, I check that the content of the degrees is close within group. For example, even though teaching and psychology lead to different occupations, they

[19]Table 2.1 describes the disciplines in each group.

are grouped together for statistical precision and because a comparison of their curricula showed several similarities. This approach is justified by the ultimate interest of this paper in understanding the role of the content of degrees.

Figure 2.4: Breakdown of fields of study taught in degrees



The figure presents groups of degrees on the vertical axis plotted against the content in each degree. Larger bubbles indicate that more credits (ECTSs) in a given group of university courses are taught in a given degree. The percentages on the diagonal refer to the time spent studying the main field of study of the degree. Off-diagonal bubbles represent the credits spent studying field of study x in degree y. A row fully describes a university degree. The left (blue) panel refers to bachelor's degrees, while the right (red) panel to master's degrees. The groups of degrees are provided by AlmaLaurea and further aggregated for statistical precision, the full description is available in appendix B.5.2. The groups of university courses are provided by MIUR and further aggregated by myself. A description of the labels is summarized in table 2.1. The unit that defines the bubble size is one ECTS (university credit).

In addition to information about disciplinary content, I also collected information about admission requirements. I do this differently for bachelor's and master's programs to account for differences in enrollment procedures.

For bachelors', I survey the admission procedures to 2296 undergraduate programs in Italy by codifying the following information: presence of a entry exam, type of exam (standardized test, multiple choice, open-end exam, knowledge assessment), number of spots, number of applicants, and application windows.[20] I use this information to construct an indicator of binding admission restrictions for each bachelor's program. Specifically, I construct a dummy for each program that is equal to 1 if the bachelor's features fewer spots than applicants in the first round of admissions. For some programs the number of applicants is not available. In these cases, I use information on the dates of opening and closing of the ap-

[20]The information on admission procedures is only widely available for the years 2018 to 2021. However, all the additional evidence that I could procure points toward high persistence in enrollment practices and admission rates.

plication phase to infer whether the selection process is competitive. Application calls that are reopened several times or that remain open well into the beginning of the program suggest that the selection process is not too stringent. Hence, in the absence of information on applicants, I classify programs with entry exams as not having binding admission restrictions in the cases where the call has been reopened or where the exam consists of a low-stakes knowledge assessment.

Admission into the master's in most instances depends on a student's ability to meet eligibility requirements in terms credits acquired during one's bachelor's. Additional criteria include bachelor's grades and interviews. Entry exams are rare, but may be in place for healthcare-related fields and psychology. Even in these cases, students must meet curricular criteria. I collect information on all eligibility requirements by surveying all public university master's programs in 2020 and 2021.[21] This information is then matched with the previously collected data on academic curricula to calculate the number of credits that must be acquired beyond those already contained in the bachelor's for each pair of undergraduate and graduate degrees. For example, a student with a bachelor's degree in economics meets all the requirements for enrollment in a master's in economics. However, she must acquire 41 additional credits to be eligible for enrollment in a master's in statistics. Conversely, any student who wants to enroll in a master's in economics must have acquired 53 credits in economics, statistics, and other social sciences. The exact number of additional credits that the student must earn will depend on the content covered in her bachelor's. When a bachelor's does not meet any eligibility criteria, the number of necessary credits is set to 180, equivalent to starting over another bachelor's degree. This is the case for access into many degrees that only admit a subset of bachelor's or single-cycle master's degrees such as law or medicine which prevent students from transferring.

The vector of exclusion restrictions Z_{ij} that regulates access into the bachelor's is built based on the previously described data on admission criteria into undergraduate programs. I build a measure of the percentage of bachelor's degrees for which the admission criterion is binding for each aggregated category of degrees as described in appendix B.5.2 and university, and merge it with the administrative data for each individual and closest public institution. There are thus ten variables, one for each group of bachelor's degrees, that measure the share of degrees within a group with a binding admission requirements in the institution closest to the individual's place of birth. As not all universities offer all groups of degrees and programs in different universities vary in their admission restrictions,

[21] Again, admission criteria are highly persistent in time such that the collected information is strongly relevant even if the years of enrollment do not match the years in which the requirements were collected.

Table 2.4: Descriptive statistics for the exclusion restriction variables Z_{ij} and Z_{im}

Variable	Mean	Std. Dev.	Min	Max
A. Z_{ij}: *Entry Exams*				
EE (AVGB)	0.492	0.183	0.100	0.883
EE (Arch.Eng.)	0.417	0.196	0	0.889
EE (Chem.Pharm.)	0.634	0.258	0	1
EE (Econ.Mgmt.)	0.453	0.377	0	1
EE (Ed.Psy.)	0.769	0.234	0.130	1
EE (Law)	0.172	0.235	0	0.759
EE (Lit.Lang.)	0.165	0.126	0.001	0.672
EE (Health)	0.939	0.068	0.791	1
EE (Pol.Soc.)	0.299	0.238	0	0.852
EE (Sci.Stat.)	0.308	0.273	0	1
B. Z_{im}: *Constrained number of credits*				
Cred. (AVGB)	60.987	17.616	0	69.874
Cred. (Arch.Eng.)	86.927	20.680	0	96.249
Cred. (Chem.Pharm.)	84.485	22.110	0	95.891
Cred. (Econ.Mgmt.)	50.539	18.686	0	58.404
Cred. (Ed.Psy.)	65.998	22.539	0	82.778
Cred. (Law)	91.487	38.442	0	114.910
Cred. (Lit.Lang.)	65.866	5.802	48.790	69.000
Cred. (Health)	146.272	33.683	0	163.571
Cred. (Pol.Soc.)	41.140	20.475	0	62.066
Cred. (Sci.Stat.)	76.325	9.041	40.040	86.584
C. Z_{im}: *Constrained number of credits (standardized)*				
Cred. (AVGB)	-0.766	0.843	-3.683	-0.341
Cred. (Arch.Eng.)	0.475	0.989	-3.683	0.921
Cred. (Chem.Pharm.)	0.358	1.058	-3.683	0.903
Cred. (Econ.Mgmt.)	-1.266	0.894	-3.683	-0.890
Cred. (Ed.Psy.)	-0.526	1.078	-3.683	0.276
Cred. (Law)	0.693	1.839	-3.683	1.813
Cred. (Lit.Lang.)	-0.533	0.278	-1.349	-0.383
Cred. (Health)	3.313	1.611	-3.683	4.141
Cred. (Pol.Soc.)	-1.715	0.979	-3.683	-0.714
Cred. (Sci.Stat.)	-0.032	0.432	-1.768	0.458

Total number of observations: 655 847; global average of constrained credits across degrees: 77.003. AVGB – Life Sciences, ArEn – Architecture and Engineering, ChPh – Chemistry and Pharmacy, EcMg – Economics and Management, EdPs – Education and Psychology, Hum – Humanities, Literature and Languages, Law – Law, Med – Medicine and Health, PlSc Political and Social Sciences, Sci – Math, Physics and Statistics.

this information will vary with the individual and the degree. Vector Z_{ij} is clearly exogenous since students cannot influence the level of applicants. Panel A in table 2.4 summarizes these ten variables, one for each bachelor's degree, that vary between 0 and 1, with 1 indicating that all degrees in a given group and institution present binding admission requirements and 0 indicating that none do. On average, the presence of binding admission requirements is lowest in humanities and highest in medicine and healthcare degrees.

The vector of exclusion restrictions Z_{im} that governs admission into master's degrees includes the measures on the differences between each undergraduate's curriculum and the enrollment requirements for all master's programs. There are ten variables, one for every master's program, that vary at the individual and program level. Panels B and C in table 2.4 summarize these ten variables, one for each master's degree, where panel B presents the average values in terms of credits, and panel C transforms the variables in panel B by standardizing them. On average, students must acquire 77 constrained credits to enter a master's program. Once again, there is substantial variation across fields of study.[22] Average admission requirements are highest for healthcare degrees and lowest for political and social sciences. I additionally include the log distance to the closest public university to instrument the choice not to enroll in a graduate program.

2.5 Returns to University Careers

This section discusses the implementation of the model outlined in section 2.3 to obtain labor market returns to combinations of undergraduate and graduate degrees. I discuss the relevant steps of the estimation procedure sequentially to highlight the information available at each stage as summarized in figure 2.1.

2.5.1 Choice of Bachelor's and Master's Degrees

Equations (2.1) and (2.2) are brought to the data sequentially even though in principle it should be possible to estimate them simultaneously through a nested logit model. However, several considerations about the data – mostly empty cell problems due to not all combinations existing and large differences in the size of degree combinations – make it more convenient to estimate the equations separately in the order presented in section 2.3.1 as multinomial logit models (equations 2.1 and 2.2).

Here and throughout this section, the vector of observed individual characteristics X_i will include high school grade, standardized at the province level to

[22]These variables are standardized in the empirical analysis to improve model fit.

account for regional differences in grading standards, high school type (humanities, scientific or other – baseline category), gender, parents' education (at least one parent with a college degree), and parents' occupation (at least one parent in a high-ranked occupation: academics, liberal professionals, entrepreneurs, executives). Summary statistics for these variables were reported in section 2.4.1. Additional controls include information on local labor markets (employment rate for 25-34 year olds in the province of birth at the time of enrollment) and an index of university quality from Censis, an independent research center, standardized to improve model fit. The battery of fixed effects Θ includes fixed effects for the year of graduation θ^{year}, macro-region θ^{geo}, and years since graduation θ^{exper}.[23] The choice set of bachelors' B is described in table 2.1 and includes ten aggregated fields of study. The variables belonging to vector Z_{ij} are the share of binding entry exams for each group of degrees in the public university closest to the student's province of birth previously described in section 2.4.2 and summarized in panel A of table 2.4.

Table 2.5 presents the results for equation (2.1). The excluded category is the choice of bachelor in humanities as it is the bachelor with the lowest average value of the instrument on the share of binding entry exams. The exclusion restrictions are jointly strongly significant with $\chi^2(90) = 46572.60$.[24] Clearly, rich substitution patterns emerge. Increasing the share of programs with binding entry exams in law and health increases the probability of enrollment in all degrees with respect to the baseline category (humanities), entry exams in other degrees have more nuanced effects. Interestingly, coefficients λ_j are positive along the diagonal for degrees in engineering, education, law, health and political sciences, such that decreasing the selectivity of these degrees decreases the relative probability of enrollment. This suggests that positive signaling through selectiveness may be an attribute of these degrees. Table B.3 in appendix B.2 additionally presents the marginal effects of coefficients λ, estimated at the mean of the right-hand variables of equation 2.1. Shifts in the share of degree programs with binding entry exams lead to substantial changes in the probability of enrolling in different degrees, along rich substitution patterns. Just like the coefficients in table 2.5, the marginal effects contained in table B.3 suggest that marginally changing the bindingness of entry exams leads to significant shifts in the probability of enrolling into different degrees at the average values of the sample. Even though on average the net shift of each

[23] I use the standard definition of macro-regions from the National Statistical Institute (ISTAT): North-East, North-West, Center, South, Islands.

[24] Each element of Z_j is also individually strongly significant with $p = 0$. AVGB: $\chi^2(9) = 3557.03$, Arc.Eng.: $\chi^2(9) = 2672.36$, Chem.Pharm.: $\chi^2(9) = 9441.17$, Econ. Mgmt.: $\chi^2(9) = 3155.64$, Educ.Psy.: $\chi^2(9) = 5385.44$, Hum.: $\chi^2(9) = 2613.46$, Law: $\chi^2(9) = 7722.88$, Health: $\chi^2(9) = 6836.68$, Pol.Soc.: $\chi^2(9) = 2787.76$, Sci.Stat.: $\chi^2(9) = 1857.02$.

instrument is close to 0, the variance of the marginal effects is highest for the entry exam variables in literature and languages and health, suggesting that students are particularly reactive to the admission policies of these degrees in their decision to enroll in higher education. I offer an additional discussion on the magnitude of the effects of the exclusion restrictions in section 2.5.1.1. Figure 2.5 shows how the model fits the data. As the estimator used to fit equation (2.1) is based on maximum likelihood, it matched group averages. To show how accurate the predictions are, I fit the model using cohorts 2007-2011 and present the average data and predictions for cohorts 2012-2014. Indeed, the model seems to match the observed choices on average quite well when I do not require matching on group averages, with differences in enrollment being less that 2 percentage points. The coefficients of equation (2.1) are eventually used to estimate the probability P_{ij} of enrolling in any bachelor's for all individuals.

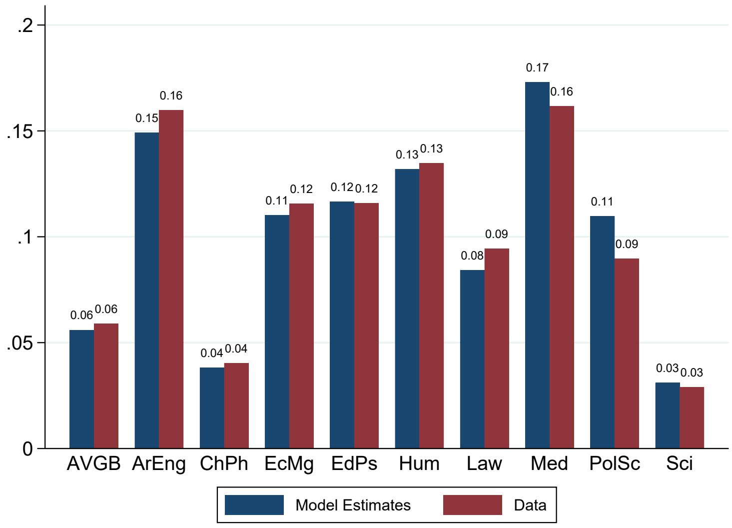
Table 2.5: Period 1 – Choice of Bachelor

VARIABLES	AVGB	Arc.Eng.	Chem.Ph.	Econ.Mg.	Ed.Psy.	Law	Health	Pol.Soc.	Sci.Stat.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Z_j: *Entry Exams*									
AVGB	-0.527***	-0.461***	-0.145**	-1.331***	0.807***	0.539***	0.526***	0.300***	0.332***
	(0.053)	(0.038)	(0.063)	(0.039)	(0.041)	(0.045)	(0.040)	(0.041)	(0.064)
Arc.Eng.	0.320***	0.637***	1.454***	-0.831***	-0.629***	-0.158***	1.428***	0.375***	1.141***
	(0.073)	(0.054)	(0.085)	(0.055)	(0.056)	(0.061)	(0.054)	(0.058)	(0.094)
Chem.Ph.	0.282***	0.501***	-0.737***	-0.015	-0.890***	-0.301***	-2.508***	-0.333***	-1.444***
	(0.047)	(0.034)	(0.056)	(0.034)	(0.036)	(0.039)	(0.035)	(0.037)	(0.057)
Econ.Mg.	-0.109***	-0.186***	-0.677***	-0.271***	-0.180***	-0.308***	-1.212***	-0.442***	-0.143***
	(0.035)	(0.026)	(0.041)	(0.026)	(0.028)	(0.031)	(0.025)	(0.028)	(0.044)
Ed.Psy.	0.329***	0.924***	-0.145***	0.325***	0.887***	-0.103***	1.853***	0.020	0.848***
	(0.043)	(0.031)	(0.050)	(0.033)	(0.032)	(0.034)	(0.032)	(0.032)	(0.055)
Law	1.848***	1.378***	1.043***	1.244***	0.813***	1.244***	1.873***	0.458***	0.736***
	(0.059)	(0.044)	(0.066)	(0.046)	(0.047)	(0.050)	(0.044)	(0.047)	(0.071)
Hum	-4.569***	-0.499***	-2.914***	0.580***	-3.077***	-2.332***	-3.874***	-2.064***	-1.658***
	(0.096)	(0.064)	(0.103)	(0.064)	(0.068)	(0.073)	(0.071)	(0.068)	(0.100)
Health	6.876***	2.693***	7.326***	4.235***	3.999***	2.987***	6.855***	1.795***	4.261***
	(0.138)	(0.102)	(0.163)	(0.103)	(0.105)	(0.113)	(0.101)	(0.108)	(0.175)
Pol.Soc.	-1.297***	-2.262***	0.893***	0.130*	-0.888***	-0.188**	0.967***	0.674***	0.373***
	(0.101)	(0.073)	(0.113)	(0.073)	(0.078)	(0.085)	(0.073)	(0.079)	(0.121)
Sci.Stat.	0.864***	-0.060	0.543***	-0.256***	1.245***	0.599***	-1.171***	0.436***	0.045
	(0.085)	(0.062)	(0.096)	(0.062)	(0.064)	(0.072)	(0.060)	(0.066)	(0.102)
X					Yes				
FE					Yes				
Observations	655,847	655,847	655,847	655,847	655,847	655,847	655,847	655,847	655,847

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Pseudo $R^2 = 0.103$.

Excluded category: humanities. Joint test of exclusion restrictions Z_j: $\chi^2(90) = 46572.60$, p-value=0. X: gender, high school grade, high school type, parent occupation, parent education, local labor market, and university quality controls. Θ: Macro-region, experience and year fixed effects.

Figure 2.5: Comparison of model and data - choice of bachelor



Model fitted on cohorts 2007-2011, predictions and data plotted for 2012-2014. Description of titles: the title refers to the previous bachelor choice on which the model is fitted. AVGB – Life Sciences, ArEng – Architecture and Engineering, ChPh – Chemistry and Pharmacy, EcMg – Economics and Management, EdPs – Education and Psychology, Hum – Humanities, Literature and Languages, Law – Law, Med – Medicine and Health, PolSc – Political and Social Sciences, Sci – Math, Physics and Statistics.

Estimating the probability of enrolling in a master's degree is slightly more cumbersome as it is conditional on the choice of bachelor's degree. I estimate ten separate multinomial logit models (equations 2.2) on the subsample of students in each bachelor's.[25] I then predict the probability of choosing any master's for all conditional choices of bachelor's $P_{im} \mid j \; \forall \; j \in B$, $i \in I$.

While the possible fields of study coincide between bachelor and master, the set of choices of master's M is different from B as it also includes the possibility of no master at all, that is, entering directly the labor market after the bachelor's. X and the fixed effects are defined as before and only vary at the individual level.[26] The omitted category is always the choice of not pursuing a master's. The choice-theoretic characterization is that not pursuing a master's is equivalent to a lack of treatment conditional on the choice of bachelor's, thus always at least the second best option. Furthermore, the option is always available. $Z_{im|j}$ is a rich set of exclusion restrictions that regulate access to the master's program and vary with the previous choice of bachelor's. It includes the standardized credit requirements for enrollment into each master's that vary at the individual and program level described in table 2.4 panel B, and log distance to the closest public university. Not

[25] Only students who are not enrolled in a single cycle degree are used to fit the model as they have to make a choice. The prediction uses the whole sample. This should not matter as the offer of single cycle degrees is plausibly exogenous to the choice and to labor market outcomes.

[26] Fixed effects for years since graduation are omitted due to collinearity with other covariates or lack of variation in certain subsamples.

all degree combinations can be estimated since some are not observed in the data (table 2.2 summarizes the available groups). Hence, only the credit requirements relevant to the possible choices are included.

Tables B.4 to B.13 in appendix B.3 present the results of these estimations. In all cases, the baseline category is to not enroll in a master's degree. Some exclusion restrictions on credit requirements may be dropped for collinearity or lack of variation within certain subgroups. For instance, this may occur if all students with the same bachelor face the same credit requirements for a given master's. Joint tests of the exclusion restrictions are presented in table 2.6 and indicate that the exclusion restrictions are valid within each conditional choice of bachelor's. Again, rich substitution patterns emerge. In all cases except one, increasing the credit requirement in the master's with the same discipline as the bachelor's decreases the probability of enrolling in that master's. Positive coefficients indicate that the probability of enrollment increases with increases in the credit requirement with respect to the choice of not enrolling in a master's. This suggests that for certain degree combinations, the probability of enrollment increases with the additional (relative) work that the student must do. Students with graduate parents are more likely to enroll in a master's degree, with very few exceptions. Gender does not seem to systematically generate sorting into more (less) quantitative fields during the master, even though it does increase the probability of enrolling in masters' in education and psychology.[27] The model fit is presented in figure 2.6 by comparing average predicted probabilities and observed enrollment. As before, equations (2.2) are estimated on cohorts 2007-2012 and the comparison between data and estimates is presented for years 2012-2014; the model seems to predict the conditional probability of enrolling in a master well.

Lastly, I estimate the probability of enrolling in any combination of degrees $P_{ijm} = P_{ij} \times [P_{m_i} \mid j]$ for all $i \in I$, $j \in B$ and $m \in M$. For the special case of students who end up in single-cycle degrees, $P_{ijm} = P_{ij}$ if $j = m$. I am left with the choice probabilities for 56 combinations of degrees.[28] On average, probabilities P_{jm} match observed treatments D_{jm}. Their difference across all degree combinations is 7.14×10^{-9}. Importantly, since P_{jm} is the product of two probabilities, the observed maximum values are strictly lower than 1, ranging from 0.012 for (Econ.Mgmt, Educ.Psy.) to 0.748 for (Healthcare, No Master), with degree combinations chosen less frequently presenting lower ranges of probabilities of enrollment. Additional

[27]Marginal effects for the exclusion restriction variables, estimated at the means of the sample are available upon request.

[28]In practice, I can only retrieve 43 returns to combinations of degrees ex post. The rationale is explained in sections 2.5.2 and 2.6. A priori, all the data from 56 combinations of degrees is used.

Table 2.6: Test of exclusion restrictions for equations (2.2)

	All Z_m		Credit Requirements			
Conditional Choice of Bachelor	D.f.	χ^2	D.f.	χ^2	Observations	Table
Agr.Vet.Geo.Bio.	25	3725.9	20	3691.69	32,494	B.3.B.4
Architecture and Engineering	20	6572.69	16	6570.72	79,817	B.3.B.5
Chemistry and Pharmacy	6	277.14	3	273.37	7,398	B.3.B.6
Economics and Management	10	14011.08	5	13977.24	75,993	B.3.B.7
P.E., Teaching and Psychology	12	10142.09	8	10106.78	62,741	B.3.B.8
Law	8	1089.64	4	1076.1	10,882	B.3.B.9
Literature and Languages	30	3083.15	24	3048.16	90,681	B.3.B.10
Healthcare and Medicine	6	861.1	3	855.77	81,883	B.3.B.11
Political and Social Sciences	36	7988.74	30	7974.45	65,798	B.3.B.12
Science and Statistics	18	1045.78	12	1040.46	20,721	B.3.B.13

Joint test of all exclusion restrictions for each conditional bachelor choice, d.f. denotes degrees of freedom. All reported χ^2 have p-values equal to 0. Z_m includes bachelor final grade (standardized), credit requirement (standardized) and distance to closest public university. Students who previously enrolled in single-cycle degrees are not used for inference.

summary statistics for the treatments D_{jm} and probabilities P_{jm} can be found in table B.14 in appendix B.3.

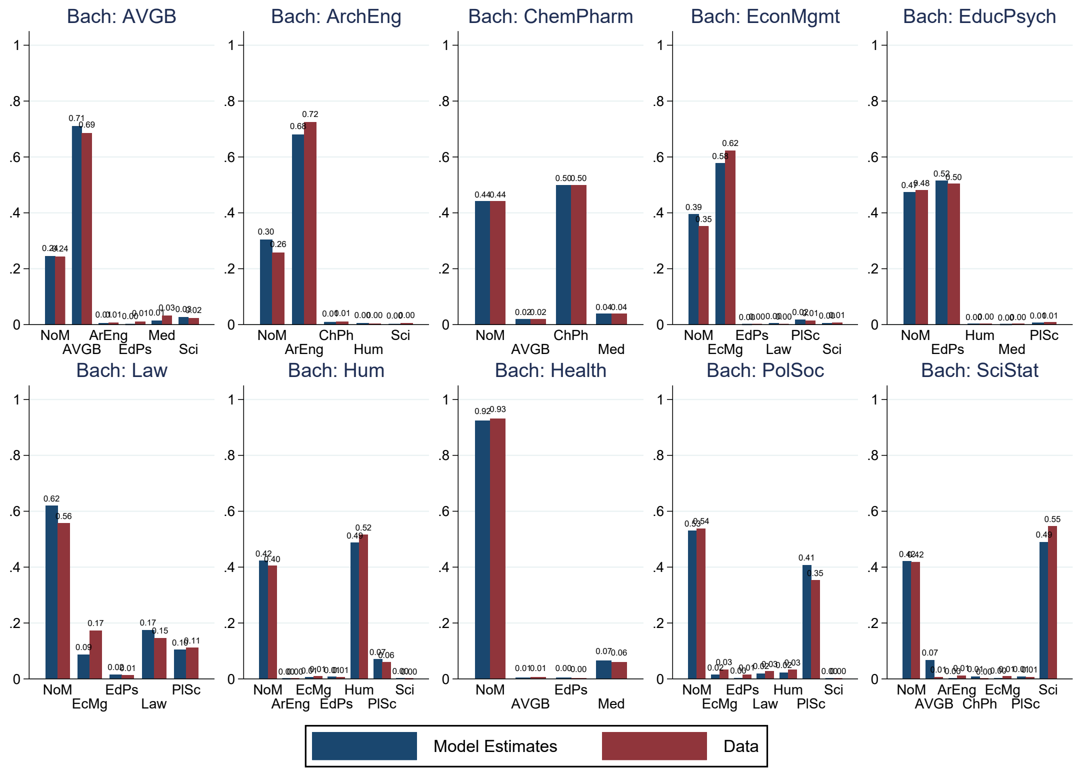
2.5.1.1 Exclusion Restrictions Z_{ij} and Simulations

I present two policy simulations that investigate different admission policies in the bachelor's to elicit how sorting at the margin responds to shifts in entry restrictions. The focus will be on entry into bachelor degrees as it leads to remarkable shifts in the student body composition. Using the choice model set up in section 2.3, I shift the values of Z_j in equation (2.1) to understand how students react to entry exams. Figure 2.5 has previously justified the appropriateness of the model to predict the distribution of students across degrees. As the available data is not appropriate to understand the labor market outcomes of individuals who did not attend college, I am unable to assess the inbound shift that might occur if admission policies were to change substantially. For these reasons, these simulations should be interpreted as shifts in enrollment at the intensive margin.

In the first simulation, all variables in Z_j are set to their minimum and new probabilities of enrollment in each degree are estimated using equation (2.1).[29] The global effect of this policy is shown in the left panel of figure 2.7 and suggests that relaxing entry barriers would increase enrollment in economics and management, humanities (literature and languages), law, and political and social sciences, while decreasing enrollment in all the other degrees. This may be rationalized by

[29] Values of Z_j are set to their observed minimum rather than 0 for all degrees because certain degrees such as healthcare have minimum values which are very high (78%), otherwise resulting in out-of-sample predictions.

Figure 2.6: Comparison of model and data - choice of master

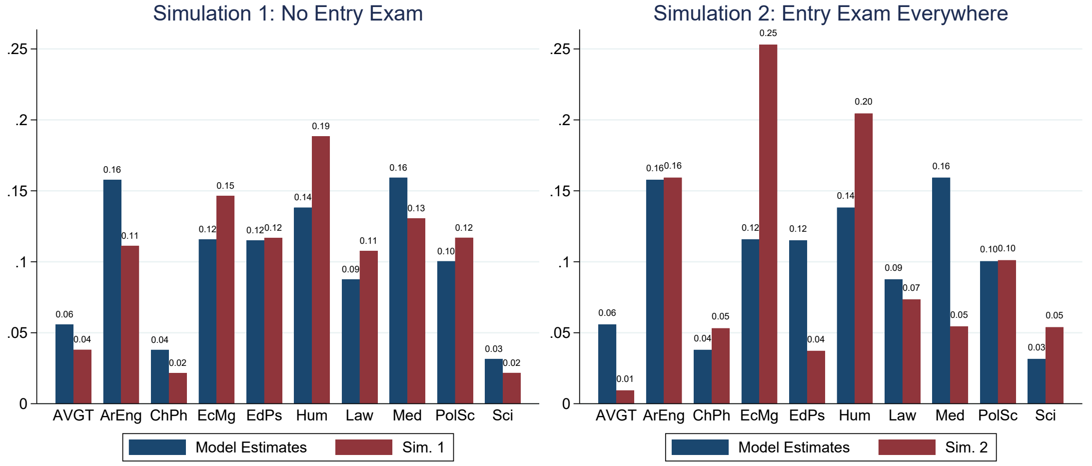


Model fitted on cohorts 2007-2011, predictions and data presented for cohorts 2012-2014. Students who enroll in single-cycle degrees (e.g. architecture, medicine, law) are not considered here as they do not make a schooling choice. The title of each histogram refers to the previous bachelor choice on which the model is fitted. Description of labels: AVGB – Agriculture, Veterinary, Geology, Biology; ArEn – Architecture and Engineering; ChPh – Chemistry and Pharmacy; EcMg – Economics and Management; EdPs – P.E., Teaching and Psychology; Law – Law; Hum – Literature and Languages; Med – Health; PlSc – Political and Social Sciences; Sci – Science and Statistics; NoM – No Master.

considering that enrollment in the former is bound by entry exams, while demand for the latter may not be determined by it. This means that if there were fewer entry exams, enrollment would increase by 35.7% in humanities (5 p.p.) and 20% (3 p.p.) in economics and management. The largest decrease would occur in engineering, with a 31% decrease in enrollment (5 p.p.). Figure B.5 in the appendix presents the results of this simulation decomposed across several individual characteristics: gender, parental occupation, education, and high school grades. While sorting into degrees varies along these dimensions, reducing entry barriers does not produce additional patterns.

An alternative simulation where entry exams are imposed everywhere is presented in the right panel of figure 2.7. Here, all variables in Z_j are set to 1 (i.e., all bachelor's programs have binding admission requirements) and new probabilities of enrollment in each degree are estimated using equation (2.1). Once again, en-

Figure 2.7: Period 1: Policy Simulations on Entry Exams



rollment in economics and humanities increases, as well as enrollment in chemistry and science. The comparison of the two simulations in figure 2.7 showcases the nonlinear substitution patterns that are possible due to the rich set of information on selective entry admissions Z_j.

Simulation 1 in figure 2.7 suggests that the existing entry exams mostly serve the purpose of managing excess demand into less quantitative fields such as economics or humanities. In fact, if students have lower preferences for quantitative studies even after controlling for rich individual characteristics (Rask, 2010; Mann and DiPrete, 2013; Fricke et al., 2018), it is not surprising that removing entry barriers does not increase enrollment into such degrees. On the other hand, simulation 2 indicates the degrees where selectiveness at the margin is positively related to enrollment. One interpretation of these results is that students derive a net benefit at the margin of increasing selectiveness in economics, humanities, chemistry, and science. I rationalize the decrease in enrollment in medicine in simulation 2 by noting that entry exams are so ubiquitously present that the signal of selectiveness is saturated at the margin. Jointly, these simulations illustrate the richness of the substitution patterns allowed by the model and suggest that settings where admission requirements are assumed to relate monotonically with preferences on enrollment do not fit real world situations.[30] Both of the proposed policies (elimination and imposition of binding entry exams) will reasonably induce reactions at the extensive margin as well as the intensive margin. Since individuals with no college are not observed, these results should not be interpreted as informative of global shifts in enrollment. However, they underline that when faced with multiple choices, several contrasting margins matter for sorting. In both cases, varying the

[30]Importantly, the assumption that the instrument P_{ijm} monotonically increases the take up of the treatment D_{ijm} stands.

values of the exclusion restrictions induces substantial shifts in enrollment across degrees. This suggests that one of the necessary conditions for identification in the reduced form presented in section 2.3 – that the exclusion restrictions be strongly relevant – is satisfied.

2.5.2 Returns to university careers

The probabilities P_{ijm} estimated in the previous section enter the reduced form equation (2.4) which is estimated with the previously described vector x and fixed effects, where the labor market outcomes of interest are log wages and employment, and jm only refers to combinations that are observed in the data.

To ensure that the coefficients α_{jm} can be interpreted as causal effects, I choose the combination of degrees (Lit.Lang., No Master) as the excluded category to proxy lack of treatment. Undergraduate degrees in humanities exhibit the lowest levels of binding entry exams and are available in 54 out of 67 public universities. Combined with "No Master", this university career serves as the most credible benchmark.

The results for the vector of coefficients β are presented in table 2.7.[31] All of the equations' standard errors are bootstrapped using full iterations of the entire model to account for the probabilities being predicted (equations (2.1)-(2.4)). For comparison, I also present OLS results where treatments D_{jm} substitute probabilities P_{jm}, thus not controlling for self-selection (equation (2.5)).

The reduced form coefficients in columns (2) and (4) of table 2.7 follow the sign and significance level of the OLS coefficients (columns 1 and 3) for almost all the main explanatory variables, where the magnitude of the effects increases. This is likely driven by the correction for endogeneity in the observed choices of university careers. Higher grades are strongly positively related to higher chances of being employed, whereby they do not improve wages (conditional on employment). Similarly, having a science high school degree improves outcomes in terms of employment, but not wages conditional on working. Surprisingly, once we control for university careers, women are more likely to be employed than men, even though they experience lower wages. This is likely due to selection on gender into different university careers.

Coefficients α cannot be interpreted as causal treatment effects without taking into account that the probabilities P_{jm} vary along a scale that is strictly smaller than one, as discussed in section 2.5.1. By rescaling the coefficient by the maximum

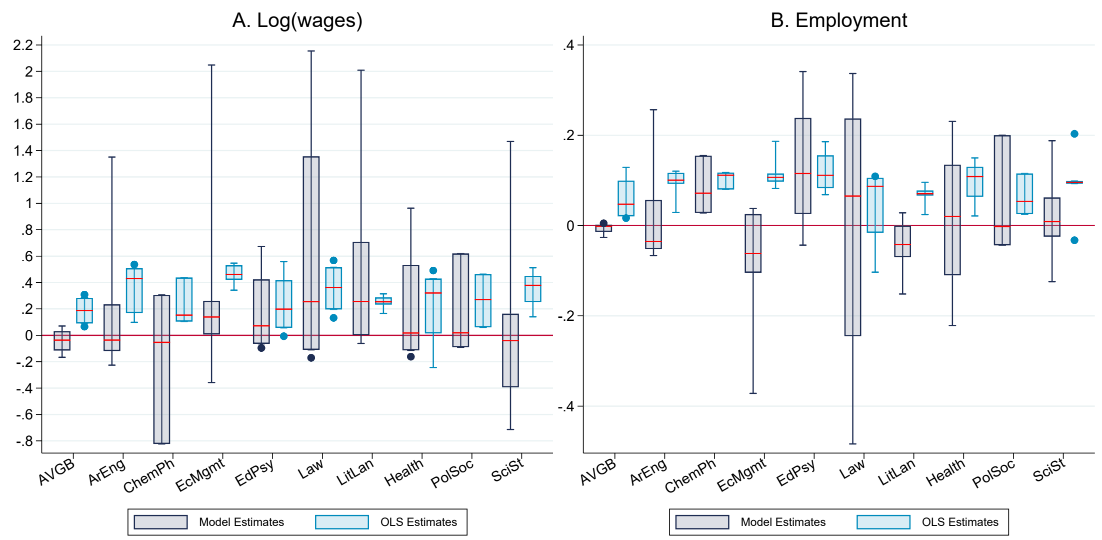
[31] Table B.1 in section B.1 reports the differences in observed characteristics X between the sample of employed and unemployed to assist the interpretation of the results on log wages conditional on employment.

observed probability of choosing a given career (j, m), the effect becomes

$$\tilde{\alpha}_{jm} = \alpha_{jm} \cdot \max_I(P_{jm}) \tag{2.7}$$

which can be interpreted as a shift in labor market outcomes induced by an increase in the probability of choosing said career from 0 to the sample's maximum, ceteris paribus.[32] In the end, I obtain 43 credible TEs for both log wages and employment.

Figure 2.8: Comparison the distributions of OLS coefficients γ_{jm} and reduced form treatment effects α_{jm}



Generalized box plots for the distribution of returns by bachelor's. Dark blue markers denote reduced form (RF) coefficients α_{jm} (2.4), light blue markers denote OLS coefficients γ_{jm} (2.5). Red markers denote medians. The baseline is (Lit.Lang., No Master).

Figure 2.8 compares the distributions of treatment effects α_{jm} and OLS coefficients γ_{jm} for university careers and both labor market outcomes and emphasizes three main findings. Notably, this comparison makes use of the strong assumptions discussed in section 2.3 that justify the IV-equivalence result of equation (2.6). OLS and reduced form results are statistically different in 84% of cases for

[32]In this setting, the causal effect of university careers (j, m) is driven by several potentially small subsamples which may display different observed characteristics, both in X and in covariate patterns of P_{jm}. Hence, when treatment effects are abnormally large (or small), it is difficult to distinguish between non-credible estimates which are not estimated precisely and credible estimates with large magnitudes due to strong self-selection. I introduce a regulating criterion to rule out treatment effects with excessive magnitudes. For employment, I ensure that all treatment effects, summed with the average predicted probability of the baseline are constrained between 0 and 1. I obtain the boundaries $\tilde{\alpha}(\text{empl}) \in [-0.62, 0.38]$ and disregard treatment effects that exhibit larger magnitudes. For log(wages), I compare the treatment effect obtained in (2.7) with the maximum (minimum) deviations from the baseline predicted in the sample. Similarly, I disregard treatment effects beyond boundaries $\tilde{\alpha}(\ln(\text{wage})) \in [-1.04, 2.27]$, in levels, this corresponds to monthly salaries between 187 and 7186 Euros. I further correct the out of sample estimated treatment effects by weighting them by the 95% percentile of P_{jm} and drop the rest.

Table 2.7: β coefficients for labor market outcomes.

	log(wage)| employed		employment	
VARIABLES	OLS	Red. Form	OLS	Red. Form
	(1)	(2)	(3)	(4)
X				
High School: grade (st.)	-0.018***	-0.937***	0.004***	3.670***
	(0.001)	(0.178)	(0.001)	(0.210)
High School: humanities	-0.079***	-1.067***	-0.032***	-0.128
	(0.003)	(0.376)	(0.002)	(0.197)
High School: science	-0.048***	-2.744***	-0.020***	15.052***
	(0.002)	(0.647)	(0.001)	(0.888)
Gender (1=female)	-0.154***	-1.956***	0.009***	3.257***
	(0.003)	(0.644)	(0.001)	(0.373)
Parents: graduate	-0.042***	-1.016***	-0.027***	4.431***
	(0.003)	(0.184)	(0.001)	(0.285)
Parents: high-ranked occup.	0.004	-0.072	0.002	1.584***
	(0.003)	(0.196)	(0.001)	(0.108)
Additional controls	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes
D_{jm}	Yes		Yes	
P_{jm}		Yes		Yes
Observations	508,242	508,242	655,847	655,847
R-squared	0.101		0.125	
Mean y	6.887	6.887	0.775	0.775

Reduced form results from equation (2.4), OLS results from equation (2.5). Columns (2) and (4) feature bootstrapped standard errors with 104 iterations. Additional controls for local labor markets and university quality.

log(wages) and in 64% of cases for employment, such that any method that does not account for self-selection into university careers is highly misleading (to compare the returns one-to-one, refer to figure B.2). Secondly, substantial variation is present when we compare the effect of university careers with the same undergraduate choice, which underscores the importance of accounting for advanced degrees in the discussion on returns to higher education. For example, log wage returns to undergraduate programs in chemistry and pharmacy vary greatly depending on the advanced degree. By plotting the distribution of the labor market returns by undergraduate choice, it is apparent that in almost all instances, the interquartile range of the conditional distribution spans positive and negative values with respect to the excluded category. Thirdly, OLS estimates more positive effects for 29 out of 43 log wage returns and 33 out of 43 returns to employment. This suggests that students self-select into degrees based on comparative advantage. Under the OLS equivalence assumptions, OLS coefficients overestimate on average the returns to university careers by 7.2pp (employment) and 0.26 log points (log wages). It also emphasizes the validity of exclusion restrictions Z_j and Z_m to partial out individual sorting. Another interpretation of these effects is thus the average returns to degree combinations enjoyed by individuals if they were randomly allocated to them. With this interpretation, it is perhaps not surprising that the average return to a career in engineering (Arch.Eng., Arch.Eng.) shifts from strictly positive when not accounting for self-selection to slightly negative when I do.

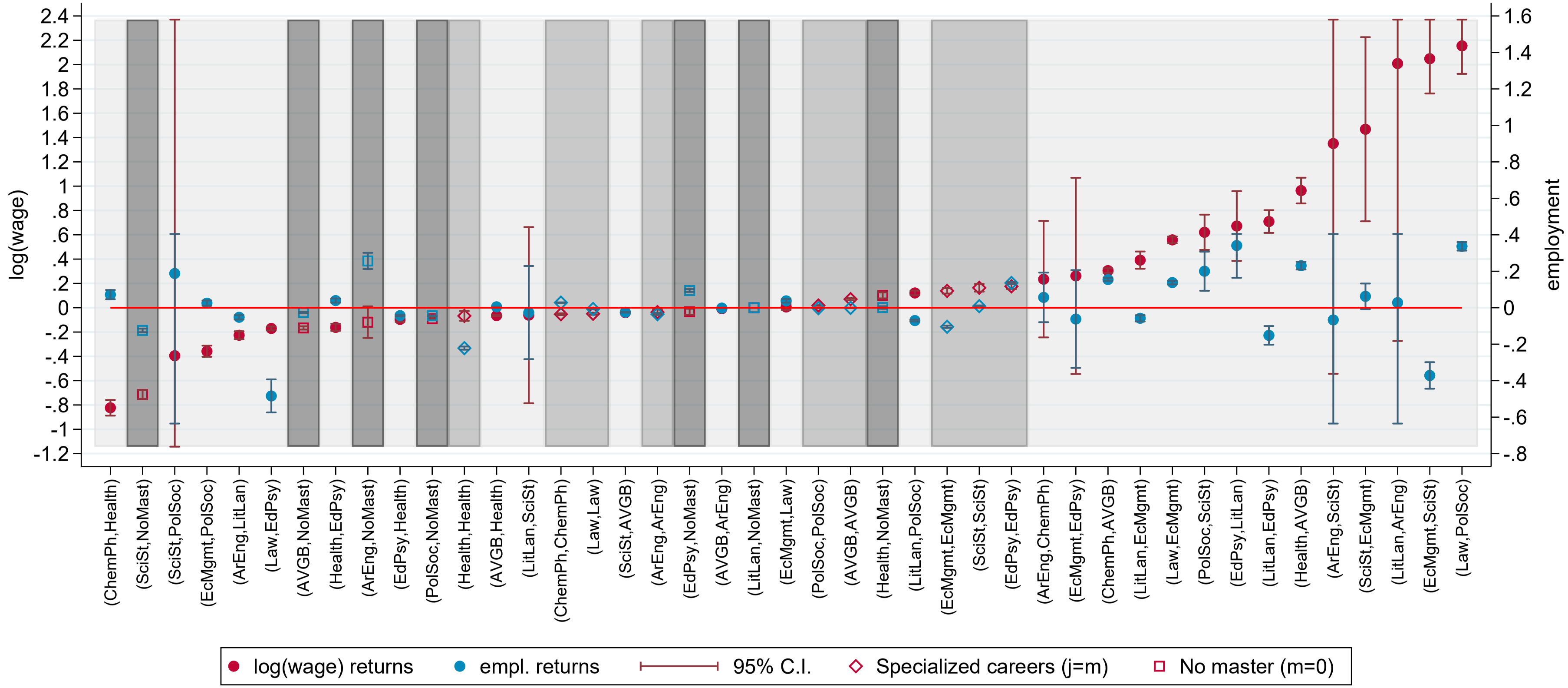
2.6 Results on Academic Curricula

Here I exploit the information on academic curricula to shed light on outcome-enhancing characteristics of university careers. I focus on how the composition of the curriculum affects returns with interest in market responses to multidisciplinary careers, quantitative courses, and the timing of degrees and courses. To facilitate the understanding of the results, I refer to careers with $j = m$ as *specialized careers*, such as (Econ.Mgmt., Econ.Mgmt.), careers with $j \neq m$ and $m \neq 0$ as *multidisciplinary careers*, for example (Econ.Mgmt., Sci.Stat.), and careers with $m = 0$ as *no master careers*, for example (Econ.Mgmt., No Master).

2.6.1 Academic Curricula and Degree Composition

Figure 2.9 directly compares the estimated returns to log wages and employment and orders careers by increasing returns to log wages. Both outcomes are significantly and positively correlated once we account for the precision of the es-

Figure 2.9: Comparison of log wage and employment returns for all careers



Comparison of log wage returns (red, left vertical axis) and returns to employment (blue, right vertical axis). Axes are centered around 0. Full-circle symbol markers indicate multidisciplinary careers $j \neq m$ (light gray shading), hollow diamond symbols indicate specialized careers $j = m$ (mid-gray shading), and hollow square symbols indicate careers with no master $m = 0$ (dark gray shading). Careers are ordered by increasing returns to log wages. Only careers for which both labor market outcomes were estimated are displayed.

timates ($\rho(\tilde{\alpha}^{\text{lnwage}}, \tilde{\alpha}^{\text{empl}}) = 0.37$, $p = 0.015$), although the relationship does not hold at the tails of the distribution of log wage returns. Especially for very high log wage returns, there seems to be a trade-off between higher pecuniary outcomes and a lower probability of employment. In the extreme case of (Ec.Mgmt., Sci.Stat.), the estimated return to log wages is 2.05 (average monthly wage of 5 784 Euros), however, the return to employment is extremely low (-0.37), resulting in a probability of employment of 24.5%. The career with the best overall outcome is (Law, Pol.Soc.), with an estimated log wage return of 2.15 (6 393 Euros) and return to employment of 0.35 (0.95 probability of employment). More generally, only 7 of the 10 careers with the highest log wage returns display positive returns to employment with respect to the excluded career (Lit.Lang., No Master). On the opposite end of the distribution, the worst overall labor market returns are associated with career (Sci.St., No Master) which features a log wage return of -0.71 (366 Euros) and return to employment of -0.12 (0.49 probability of employment).[33] These results might be partially driven by different timelines that affect entry into the profession. The pathway to employment might be more complicated for individuals with peculiar university careers, for example, because of additional requirements regarding certification, training, or difficulty in building a client base. Certain careers require long apprenticeship periods after graduation (teachers, lawyers, doctors). In other instances, differences between wages and employment may reflect the riskiness of the career, whereby few individuals reap substantial benefits (creative careers, policy). Similarly, low-earning careers with relatively high levels of employment might reflect lower riskiness of the career, which is often the case for careers with no master.[34] These low-earning careers also exhibit differences in the sign of the two labor market returns, with 5 out of the 10 lowest earning careers displaying positive returns to employment with respect to the excluded career. Figure B.4 in the appendix concentrates on careers with no master's and specialized careers that mostly populate the central part of figure 2.9. By considering the returns as a whole, I note that the returns to combinations with no master are ranked towards the bottom of the distribution of wage returns (dark gray shading), suggesting that in most instances there is a premium to having a master's degree. Specialized careers are bunched towards the middle of the distribution in mid-gray shading with sensible rankings (science

[33]In terms of employment, the worst performing university career is (Law, Ed.Psy.), with an employment coefficient of -0.48 (13.5 probability of employment on average) and a -0.17 log wage coefficient (628 Euros).

[34]The magnitude of the estimates is obtained by comparing the returns to the predicted outcomes for the excluded career (Lit.Lang., No Master) at sample averages of the observed characteristics. Value in levels (Euros) of log wage return $\tilde{\alpha}_{jm}^{\text{lnwage}}$ is $\exp(6.614 + \tilde{\alpha}_{jm}^{\text{lnwage}})$, probability of employment for return to employment $\tilde{\alpha}_{jm}^{\text{empl}}$ is $0.615 + \tilde{\alpha}_{jm}^{\text{empl}}$.

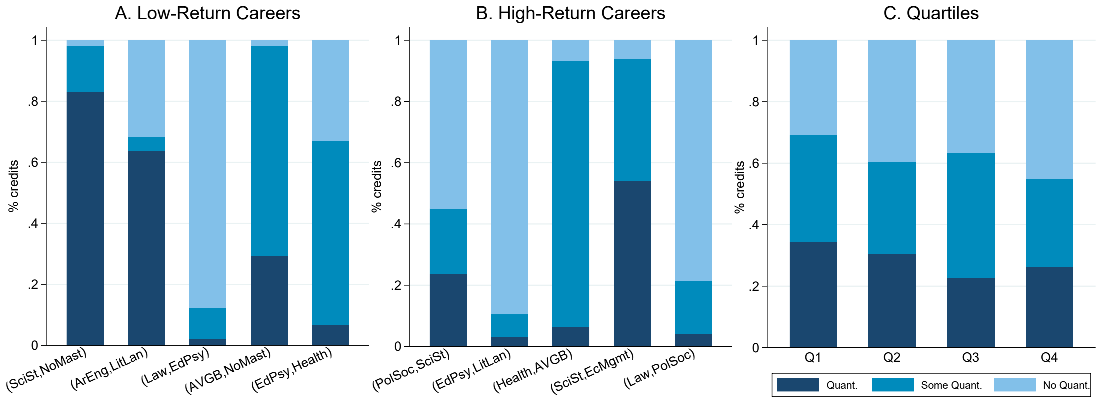
ranks better than economics which ranks better than law), while the top of the distribution is exclusively populated by multidisciplinary careers (light gray shading).[35] Out of 43 estimated returns to careers, the 14 highest log wage returns are all multidisciplinary careers with $j \neq m$ and $m \neq 0$ (top third of the distribution), while the 10 lowest log wage returns are associated with no master careers in 3 cases and multidisciplinary careers in the other 7. These findings suggest that enrolling in a multidisciplinary career can substantially boost labor market outcomes if chosen well. Even though career (Econ.Mgmt., Econ.Mgmt.) yields the third highest log wage returns among specialized careers ($\tilde{\alpha}^{\text{lnwage}}_{(\text{EcMg},\text{EcMg})} = 0.14$), returns can be up to fourteen times higher if combined with other degrees such as (Econ.Mgmt., Educ.Psy.), (Law, Econ.Mgmt.), or (Econ.Mgmt., Sci.Stat.), yielding $\tilde{\alpha}^{\text{lnwage}}_{(\text{EcMg},\text{EdPs})} = 0.26$, $\tilde{\alpha}^{\text{lnwage}}_{(\text{Law},\text{EcMg})} = 0.56$, $\tilde{\alpha}^{\text{lnwage}}_{(\text{EcMg},\text{Sci})} = 2.05$, respectively. At the same time, multidisciplinarity can lead to drastically lower returns. For example, log wage returns to (Econ.Mgmt., Pol.Soc.) are equal to -0.36, or 1.4 times lower than the specialized career.

While figure 2.9 highlights the importance of the joint choice of bachelor's and master's beyond undergraduate majors, it does not reveal which characteristics of the careers are informative about outcomes. I investigate the composition of the curriculum of the best- and worst-performing careers to elicit any patterns in the type of knowledge that is covered. In order to avoid considerations on the trade off between employment and wages, I focus on the five best-performing careers – compared to the benchmark – which display the highest log wage returns as well as positive returns to employment. Similarly, the five worst-performing careers are selected such that they display negative returns to both outcomes.

Panels A and B of figure 2.10 present the academic curricula of the selected high- and low-performing careers. The curricula are summarized as the share of credits in courses with different levels of quantitative content. Following the agreement among scholars in the categorization of STEM disciplines (table B.2), I group university courses according to their quantitative content. Quantitative courses include science and statistics, architecture and engineering, and chemistry and pharmacy. These are the fields of study that most scholars agree can be defined as STEM. Courses with some quantitative component include life sciences (agriculture, veterinary, geology and biology), economics and management, and healthcare. These are more technical fields of study over which researchers disagree

[35] Indeed, some specialized careers result in surprising results: the best-ranking specialized career is education and psychology, while the worst one is healthcare. The ones reported as sensible rankings remain stable throughout versions of this paper, while the ones cited in this note change (previous versions of this paper are available upon request). Furthermore, healthcare requires extensive training after the degree such that potential long term returns are not captured in this framework, and overall the returns to specialized careers are close to each other in magnitude, leading to variations in rankings even without substantial changes in the estimated.

Figure 2.10: Comparison of academic curricula and log wage returns for selected careers

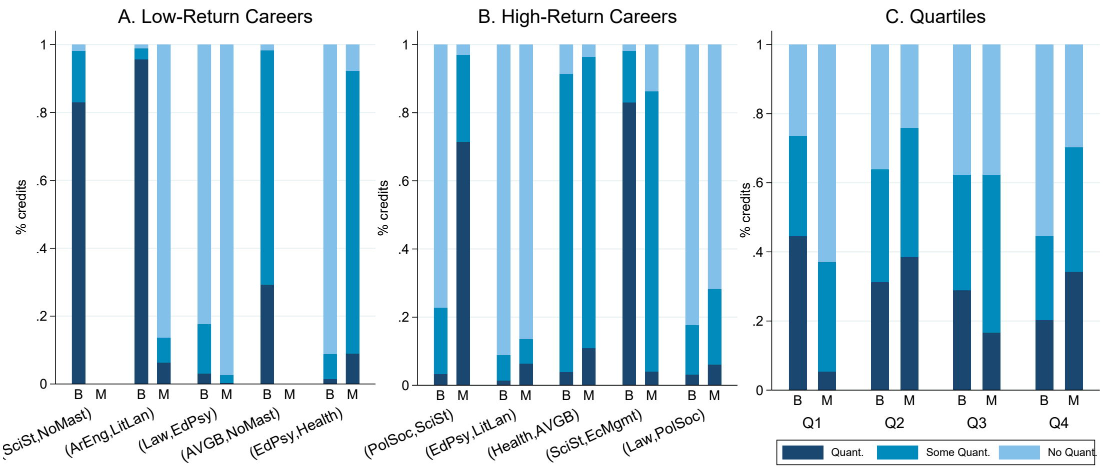


Quantitative courses (dark blue): science and statistics, architecture and engineering, and chemistry and pharmacy. Some quantitative (technical) courses (blue): life sciences (agriculture, veterinary, geology and biology), economics and management, and healthcare. Non-quantitative courses (light blue): education and psychology, law, humanities (literature and languages), and political and social sciences. The total percentage of credits in each grouping is plotted on the vertical axis. The order of degrees follows the ranking of log-wage returns, increasing from left to right within each panel.

on whether they should belong to STEM education (I will alternatively refer to these disciplines as "technical"). Non-quantitative courses include education and psychology, law, humanities (literature and languages), and political and other social sciences. Most scholars agree that these fields of study do not fit the STEM definition. The ordering along the horizontal axis reflects increasing log wage returns. Figure 2.10 shows that quantitativeness alone does not explain the higher returns of certain careers. In fact, careers with high shares of credits in quantitative courses are represented both among the worst- and best-performing careers. Panel C presents the average composition of careers by quartiles of the distribution of log wage returns. This ensures that the lack of relationship between the share of quantitative credits and returns is not driven by the choice of low- and high-return careers. Indeed, the share of quantitative courses displays a slight U-shape relationship with log wage returns. The overall share of non-quantitative courses tends to increase along the distribution of log wage returns. Panel A in figure 2.12 presents the same decomposition of academic curricula for the distribution of returns to employment. Increasing the share of quantitative courses only improves outcomes up to the third quartile, whereby the fourth quartile has the highest share of non-quantitative courses.

I report the curriculum composition for the same groups of careers separately for the bachelor's and the master's degrees to elicit patterns in the timing of courses in figure 2.11. The most striking difference between low- and high-earning degrees in terms of curriculum that emerges once courses are plotted separately by bachelor's and master's is that degrees with low returns have a low share of technical courses in the bachelor's (panel A and quartile 1 of panel C). Conversely, high-return careers have a low share of non-quantitative credits in the master's (panel B and quartile 4 of panel C). Once again, a U-shaped relationship between the share of quantitative courses and log wage returns emerges, reiterating that quantitativeness alone does not explain higher returns, even when I account for timing. Panel B in figure 2.12 presents the same decomposition of academic curricula for the distribution of returns to employment. In this case, high-performing degrees spend more time in more general type of courses in the bachelor's (non-quantitative and quantitative) while they invest substantially more in technical courses in the master's (quartile 4). These results are consistent with the paradigms of education, whereby more general education should be approached earlier and more vocational education later.[36]

Figure 2.11: Comparison of academic curricula for selected careers

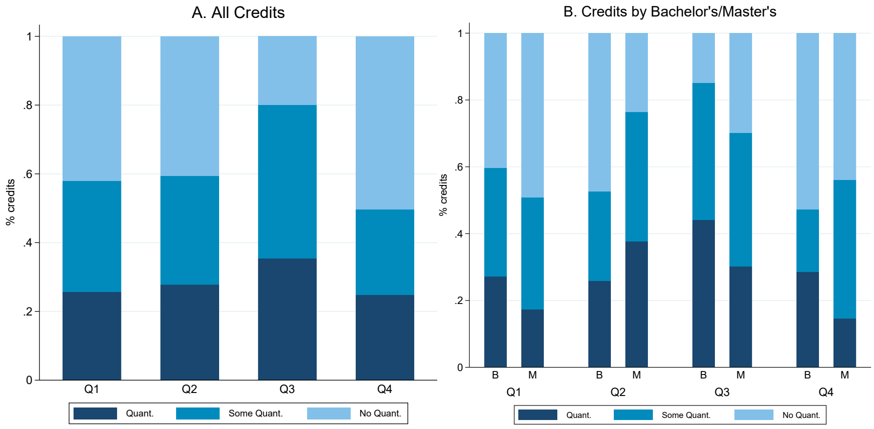


Quantitative courses (dark blue): science and statistics, architecture and engineering, and chemistry and pharmacy. Some quantitative (technical) courses (blue): life sciences (agriculture, veterinary, geology and biology), economics and management, and healthcare. Non-quantitative courses (light blue): education and psychology, law, humanities (literature and languages), and political and social sciences. The total percentage of credits in each grouping is plotted on the vertical axis. The order of degrees follows the ranking of log-wage returns, increasing from left to right within each panel. Column labels B and M denote bachelor's and master's, respectively.

To further understand how the timing of degrees affects returns, I compare the returns and career composition for symmetric multidisciplinary careers, that is, given two fields of study x and y, the returns to career (x, y) compared with career (y, x). Complete returns for both sets of outcomes are available for seven pairs of reciprocal degrees: (AVGB, Health), (Econ. Mgmt., Law), (Educ. Psyc., Health), (Pol. Soc., Sci. Stat.), (Ec.Mgmt., Sci.Stat.), (Arch.Eng., Lit.Lang.),

[36] Neal (2018) on optimal life cycle investments in skills, "*learn to learn, learn to earn, earn*" (Appendix I.5).

Figure 2.12: Comparison of academic curricula along quartiles of the distribution of employment



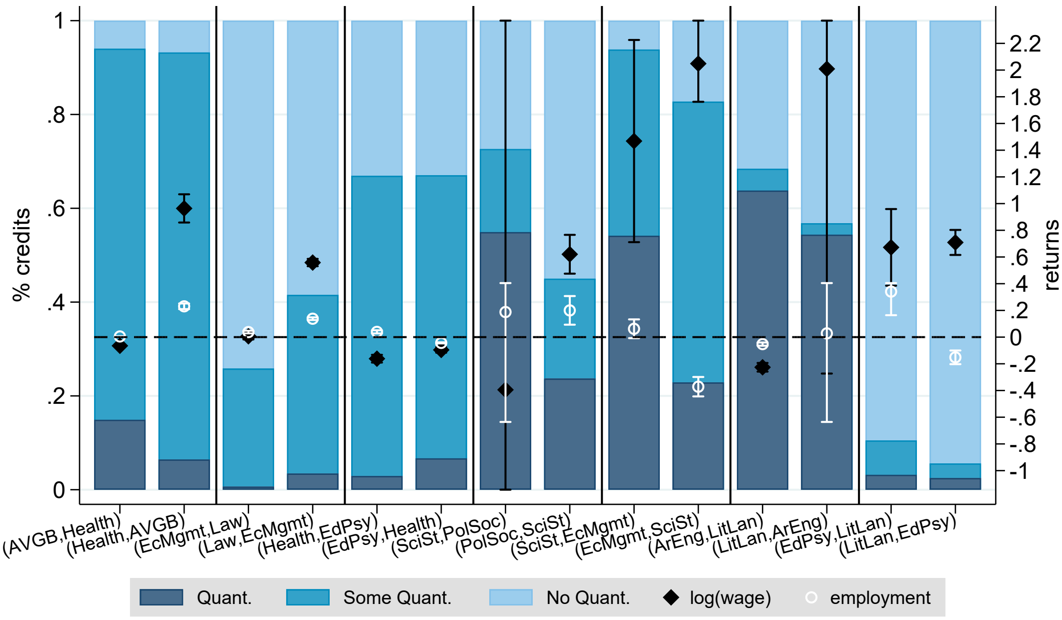
Quantitative courses (dark blue): science and statistics, architecture and engineering, and chemistry and pharmacy. Some quantitative (technical) courses (blue): life sciences (agriculture, veterinary, geology and biology), economics and management, and healthcare. Non-quantitative courses (light blue): education and psychology, law, humanities (literature and languages), and political and social sciences. The total percentage of credits in each grouping is plotted on the vertical axis. The order of degrees follows the ranking of returns to employment, increasing from left to right within each panel. Panel B further decomposes by bachelor's (B) and master's (M).

(Ed.Psy., Lit.Lang.), and the reciprocals of these groups. Figure 2.13 presents the composition of these careers by degrees and the labor market returns, where each reciprocal is ordered such that the more quantitative group of the two is studied in the master's. Even though the composition of symmetric careers is comparable, the log wage returns vary substantially. In particular, log wage returns are higher when the more quantitative of the degrees is studied later, consistent with the findings of figure 2.11.[37] This trend in log wage returns is only partially carried across returns to employment.

The analysis on the composition of curricula suggests that multidisciplinary careers can substantially increase or decrease labor market returns. While there is no clear recipe for a successful university career, several clues guide indicate best practices in the design of university programs. Quantitative courses are connected to log wage returns by a U-shaped relationship, whereby both low- and high-performing careers display relatively high shares of quantitative courses, and returns to employment increase with the share of quantitative courses only up to the third quartile. The timing of courses matters with higher shares of non-quantitative courses in the master's being related to lower returns. All high earning careers are characterized by relatively more general education early on (especially non-quantitative), and more technical courses in the master's. This is consistent

[37]A degree in Agr.Vet.Geo.Bio. contains more quantitative courses (e.g. math, chemistry) than a degree in Health. Similarly, a degree in Health contains more quantitative courses than a degree in Education and Psychology and so on.

Figure 2.13: Differences in returns for symmetric careers



Symmetric careers are grouped next to eachother. The share of courses by quantitativeness are plotted on the left vertical axis, while the estimated returns to log wages (black diamonds) and employment (white circles) follow the right vertical axis.

with the comparison of symmetric multidisciplinary careers. It suggests that returns are different even when the overall structure of the curriculum is similar, with the returns being higher for careers with the most quantitative and technical degree studied later, even if globally it may result in less time spent in these subjects.

Even though these results should not be regarded as conclusive insights on the role of timing, multidisciplinarity and quantitativeness on labor market outcomes, they do suggest that these characteristics strongly affect outcomes and call for a deeper understanding of synergies across courses. When optimally designing a degree, additional constraints on total credits, substitution patterns and complementarities between courses should be considered, as well as measuring skill acquisition at university and skill use during the job, which are not observed in this setting. While this project does not allow for an in-depth discussion of how to increase the labor market returns of existing university careers, it does suggest that the combination of quantitative and technical courses is important for labor market outcomes, that well-thought multidisciplinary careers can lead to impressive labor market outcomes, and that timing of courses matters. In particular, it does seem that specializing in quantitative degrees in graduate school is positively associated with outcomes. Indeed, the signaling component of the degree might play a role in these results, so further research is needed to corroborate the role of timing.

2.7 Conclusions

This article proposes a new method to causally estimate the returns to many combinations of bachelor's and master's degrees. It then leverages information on the course content of programs to investigate how multidisciplinarity, quantitativeness, and timing affect returns. I find that considering the joint choices of bachelor's and master's degrees is crucial to truthfully evaluate the effect of higher education on outcomes. Combining degrees in different fields can boost labor market returns, although there is no unique pattern of quantitative course content and timing that explains the success of certain careers. In fact, a U-shaped relationship between labor market returns and the share of quantitative courses emerges. The breakdown of this relationship by bachelor's and master's suggests that successful careers have little non-quantitative education in the master's, but a deeper understanding of the complementarities between courses acquired early and late in the career is necessary. Finally, policy simulations that remove entry barriers in the bachelor's suggest that students have preferences for non-quantitative degrees.

These results suggest that policies that incentivize enrollment in STEM education without considering nonlinearities in the relationship between quantitative education and outcomes might not benefit students. Furthermore, policies that incentivize STEM education through a reduction in entry barriers might be ineffective due to individual preferences, and unable to affect the composition of the student body, for example by increasing female enrollment. The results point to the importance of covering multiple disciplines throughout higher education with surprising effects on wages, challenging the prejudice that extreme specialization is profitable. This suggests that policies that ease switching from one field to another may be extremely beneficial to students. Caution is nonetheless advised, as certain combinations that encompass multiple fields can be nefarious.

One limitation of this setup is that it does not consider students' reactions to enrollment policies at the extensive margin. Indeed, the negative effects of policies that incentivize STEM enrollment through reductions in entry barriers might be attenuated if they generate a sufficient influx of students who would otherwise not obtain a degree. Furthermore, it does not incorporate the signaling component of degrees. If employers only observe the highest level of education (as assumed by Altonji (1993)), master's degrees might be weighted disproportionately by the employer, thus partially explaining the results on timing. Lastly, while the policy simulations hint at preferences towards non-quantitative studies outweighting quantitative preferences, the model does not isolate the effect. Non-pecuniary returns not captured by the model might explain some features of the sorting, in which case policies that affect enrollment could have a greater impact if they can incorporate these amenities.

This paper reveals two potential venues for future research that would improve our understanding of how knowledge acquired at university plays into the labor market. Skills are acquired during university and vary across fields, but I do not observe them in this setting. In particular, the results on the content of degrees signal the importance of the time spent in technical courses, such as medicine or management, which typically involve the acquisition of practical knowledge. We can speculate that part of the commonly observed success of STEM can be explained by the successful integration of quantitative and technical education that interplay with skills. Similarly, the concept of quantitativeness remains elusive and we can expect high returns to specific types of quantitative education. Understanding how specialized knowledge in quantitative fields and how the mathematical language spills over into different courses – for example through enhanced problem solving ability – might be critical to optimally designing university degrees.

CHAPTER 3

Credit and Voting

Joint with Giacomo De Giorgi and Jérémy Laurent-Lucchetti

There is a tight connection between credit access and voting. We show that uncertainty in access to credit pushes voters toward more conservative candidates in US elections. Using a 1% sample of the US population with valid credit reports, we relate access to credit to voting outcomes in all county-by-congressional districts over the period 2004-2016. Specifically, we construct exogenous measures of uncertainty to credit access, i.e. credit score values around which individual total credit amount jumps the most (e.g. around which uncertainty on access to credit is the highest). We then show that a 10pp increase in the share of marginal voters located just around these thresholds increases republican votes by 2.7pp, and reduces that of democrats by 2.6pp. Furthermore, winning candidates in more uncertain constituencies tend to follow a more conservative rhetoric.

3.1 Introduction

The opportunity for all US citizens to climb the social ladder and pursue the "American dream" is a central tenet of the US social contract. This requires that upward economic mobility remains accessible and widely visible. Access to credit and home ownership are considered pillars of socioeconomic mobility as they allow for wealth accumulation, e.g. by building equity and reducing housing costs. Recent research highlights that these advantages further translate into higher inter-generational income mobility (Herkenhoff et al., 2021). Despite a majority of Americans aspiring to home ownership, they face several common hurdles, such as difficulties in putting together a down payment and the ability to access credit. This topic is a highly divisive issue in the US political landscape.

Even though home ownership appears to be a shared value across party lines, it is generally purported that Republican positions in terms of regulation and government presence in mortgage markets are much laxer than Democratic positions (Hall and Yoder, 2022). For example, according to recent surveys, two-thirds of both Republicans and Democrats agree that owning a home is necessary to live the American Dream and 73% of Republicans and Democrats believe that owning a home increases a person's standing in the local community.[1] Republicans typically favor easier access to personal credit and lower banking regulations, at the cost of higher exposure to downside risk and higher individual liability. For example, the republican platform for the 2020 election was pushing toward incentives for accessing mortgages, regulatory downsizes, and minimizing the federal role in zoning decisions; while the Democratic platform promised robust investments in affordable housing production and rental assistance through the development of the Federal Housing Administration (FHA).[2]

Historically, both parties took clear opposite positions on the regulation of credits and mortgages. For example, the Glass-Steagall Act which separated commercial and investment banking and increased banking regulation in 1933 was written by two Democratic representatives, while the repeal of the Glass-Steagall Act in 1999 by The Gramm-Leach-Bliley Act was written by three Republicans. Similarly, the Dodd-Frank Act of 2010 which tightened financial regulation and increased consumer protection was written by Democrats and received strong republican opposition. This latest Act was partly dismantled under the Trump administration with the objective to ease mortgage loan data reporting requirements for the overwhelming majority of banks (as it reduced the number of banks subject to heightened regulatory scrutiny). In view of these examples, one can assume that access to mortgages and credit is strongly linked to political behavior as it is a clearly polarized issue over a central topic for most US citizens.

In this paper, we uncover such a link by directly connecting the ability to borrow to voting behavior. Specifically, we assess the effect of the proportion of individuals around salient credit score thresholds on the share of voters for each party at the county-by-congressional-district level. We show that voters tend to favor Republican candidates in districts that contain more voters around specific credit thresholds, where the probability and size of credit differ substantially within few score points, i.e. below such thresholds access to credit tightens substantially. We interpret this result as arising from higher uncertainty in access and quantity of credit that favors Republican candidates who typically run on platforms of deregulation and easier access to credit. The credit uncertainty and

[1]See `https://www.prnewswire.com/news-releases/homeownership-is-a-shared-value-across-party-lines-300553639.html`

[2]See `https://nlihc.org/resource/democratic-party-and-republican-party-platforms-address-affordable-housing`.

its impact on access to home ownership might also favor more conservative candidates through cultural channels as voters tend to be more attracted to conservative rhetoric in times of economic hardship, for example by blaming minorities or foreigners for limiting social mobility (see Funke et al. (2016) or Algan et al. (2017)), or through its effect on the fear of loss of status (Mutz (2018) and Guriev and Papaioannou (2022)). The recent literature introducing social identity in voting models (Bonomi et al. (2021) or Grossman and Helpman (2020)) shows how adverse economic shocks may generate both a behavioral response that strengthens one's identification with a specific social group – e.g. the white working class – and material interests. Higher economic uncertainty may therefore increase the political relevance of racial and ethnic identities among voters, along with support for culture-based politics, where nationalist and tribal sentiments are salient.

Our empirical approach relies on several data sources. First, we use proprietary data on a random 1% sample of the US population with a valid credit score in 2010, and yearly reports (drawn on June 30th of each year) for the same population from 2004 to 2016 (similar data are described in Lee and van der Klaauw (2010)). This extensive database of about 2 million individuals, observed for 13 years, allows us to construct new data-driven credit score thresholds below which credit tightens substantially. Specifically, we determine the credit score values around which individual total credit amount (including mortgages) jumps the most, for each local labor market (commuting zone) and year over our period of analysis.[3] These thresholds are highly relevant as the total amount of credit increases by 20,000USD on average when crossing the threshold by 1 point (where the average total amount of credit is around 100,000USD over the whole sample). We then compute the share of individuals around these thresholds – we experiment with bandwidths of 5 to 25 credit score points, at 5-point increments – for each of the 4931 US county-by-congressional districts. We then exploit the quasi-exogenous variation in these shares, across space and time, to assess their effect on US Congressional elections and ideological position. We use data on US elections at this same level of aggregation from Leip (2017), focusing on US Congressional candidates to exploit the considerable number of candidates in these races. The information on their ideological positioning on the liberal-conservative spectrum is elaborated through DW-NOMINATE scores (Poole and Rosenthal, 1985, 1991). A wealth

[3] We focus on thresholds in the credit score range of 560-650 as it is usually perceived to be a relevant range for uncertain mortgage and credit access, for example, 30.1% of new mortgages were originated in that range up to 2008, after which that figure dropped to 17.5% between 2008 and 2016 (`https://www.newyorkfed.org/medialibrary/interactives/householdcredit/data/pdf/HHDC_2015Q3`). This range encompasses subprime and near prime individuals `https://www.experian.com/assets/consumer-information/product-sheets/vantagescore-3.pdf`. Prime individuals are inframarginal with respect to credit availability, as their approval odds are quite high.

of ancillary datasets complete the analysis: demographic information from the US Census Bureau at the ZIP Code Tabulation Areas (ZCTA) level as well as geographic relationship files to build crosswalks between different geographic units, and exposure to international trade in each local labor market (Autor et al., 2020).

Our identification strategy exploits the county-by-congressional district's variation in the share of marginal individuals. Our empirical specifications account for time-invariant local characteristics and state-by-year variation and also control for population traits at the same level of geographical variation (such as race and gender). We also account for the (instrumented) share of imports from China in each district, as it is one of the major economic drivers for republican votes identified in the recent literature (Autor et al., 2020). The main identifying assumption is therefore that the variation in the number of people clustered around a specific credit threshold is an exogenous measure of credit access. Our identification argument is firstly based on the objective fact that individuals do not know where these thresholds are and therefore cannot easily manipulate their credit score to be on either side. Banks do not explicitly tell their customers where those thresholds are, and marginal customers would not have full control of their score to the 1-5 points level. If someone has a credit score within 5 points of an unknown threshold it would be very hard to believe that they can exactly infer the threshold and push their score just above it. Our argument is also strengthened by the existing evidence provided by Agarwal et al. (2017). Furthermore, our threshold computation design implies that the share of individuals around a specific salient threshold is not generated mechanically by a simple distributional shift of credit score (as thresholds are computed for each year and each commuting zone). This granularity in the data allows us to infer that credit access uncertainty is the driving factor of differential voting behavior and not, e.g., a decrease in all credit scores in the area or an increase in some part of the distribution of credit score that would generate an income effect. We discuss the methods used to infer the thresholds in detail in section 3.4.2.

Leveraging the variation in the share of potential voters around the estimated thresholds we find that a 10pp increase in the share of marginal (potential) voters increases the republican votes by 2.7pp, and reduces that of democrats by 2.6pp. These are sizeable, and robust effects (see Section 3.5), at the margin they would determine winners. Further, we show that the results are not driven by the share of potential voters below vs. above the thresholds, it appears that it is the uncertainty of being around it that matters. Using the DW-Nominate score – computing the ideology of candidates on social and economic issue based on their roll-call – we also show a substantial conservative shift for elected candidates in areas with a higher share of voters facing credit-uncertainty (especially democrat candidates). We interpret this result as consistent with the 'cultural channel' as candidates of both parties tend to become more conservative (including on social dimensions) in areas with a higher share of people experiencing credit uncertainty.

Our paper is at the crossing of several literature. First, we contribute to the literature linking economic uncertainty or hardships with a drift toward conservative voting. A large set of papers study the impact of trade (i.e. Colantone and Stanig (2018), Caselli et al. (2020) or Autor et al. (2020)) and economic crisis (Funke et al. (2016), Guiso et al. (2017)) in explaining the rise of conservative or populist candidates or populist platforms (Becker et al. (2017)). A recent paper (Kara and Yook (2022)) also shows that banks reduce the supply of mortgage loans when policy uncertainty increases in their headquarters states.[4] Finally, Mian et al. (2010) show that representatives whose constituents experience a sharp increase in mortgage defaults are more likely to support the Foreclosure Prevention Act (preventing manipulative foreclosure practices disproportionately harmful to communities of color). Another strand of the literature focuses on linking labor market conditions to conservative voting (Algan et al. (2017)) and austerity (Guriev and Papaioannou (2022)). We contribute to this literature by highlighting the unique role of credit uncertainty on voting behavior. Credit access is a pillar of social mobility and we highlight its role in voting behavior. As it is usually easier to regulate credit markets than manipulate macroeconomic conditions, this has important implications for policymaking.

Second, we speak to the literature that highlights the identity, status, and cultural roots of modern populism. People care deeply about non-monetary factors such as identity, fairness, and status (see Bénabou and Tirole (2006); Enke (2020); Guriev and Papaioannou (2022), among many others). Some recent literature shows how these sentiments affect voting. Rodrik and Mukand (2018) highlight the role of "identity politics" which focuses on changing voters' perceptions of who they are. By discussing pride and victimhood, identity politicians create an "in-group" sentiment that helps explain why low-income voters may support a right-wing politician who advocates less redistribution. Along similar lines, Enke (2020) provides evidence that the rise of populism is related to the gradual shift of Americans' moral values away from universalist and toward communal ones while Mutz (2018) argues that Trump supporters were mostly driven by the threat to their status within the society. Autor et al. (2020) also support the culture view: the China shock boosts Trump and conservative Republicans' support only in counties with (non-Hispanic) White majorities. Colantone and Stanig (2018) produce similar evidence for Europe showing that regions hit hard

[4] Notice that our design is not subject to the potential reverse causality issue implied by this result as we do not exploit the variation of the credit threshold per se.

by Chinese imports are less supportive of democratic institutions and less likely to hold liberal values. We contribute to this literature by showing that credit access uncertainty impacts voting behavior potentially through its effect on social mobility and through the status of home ownership.

Finally, we contribute to the literature identifying the various socioeconomic impacts of credit access: (see Herkenhoff et al. (2021) on the impact of consumer credit access on self-employment and entrepreneurship; Herkenhoff et al. (2016) on job finding; and for an extensive discussion on how credit access affects human capital investment and mobility see Heckman and Mosso (2014)).

The rest of the paper is organized as follows. Section 3.2 discusses the conceptual framework; Section 3.3 presents the data used for the analysis. Section 3.4.2 details the process of identifying thresholds for credit access; Section 3.5 presents the main results; and Section 3.6 concludes.

3.2 Conceptual Framework

Climbing the ladder of American society starts with the ability to pursue one's dream: rising through the ranks as a property owner, small entrepreneur, and the like. The inability to do so because of the lack of capital can push voters towards political parties with a higher propensity to be pro-credit access, pro-business, and with laxer credit regulations. It is generally understood that Republican positions in terms of regulation and government presence are much laxer than Democratic positions. And as such in favor of easier access to personal and business credit.

We draw a direct connection between credit uncertainty, i.e. the random nature of being approved for credit for marginal individuals, and political choices by the electorate. In particular, we conjecture that voters who are uncertain about their ability to access credit will disproportionately favor republican candidates as they see these candidates' positions as more in line with their needs. Suppose an individual is considering purchasing her first home, a fundamental step towards the dream, and has the need for a mortgage which she would request at the local bank. If her credit score puts her in some low and uncertain probability of obtaining such loan, she would support political positions which would increase her approval rates.

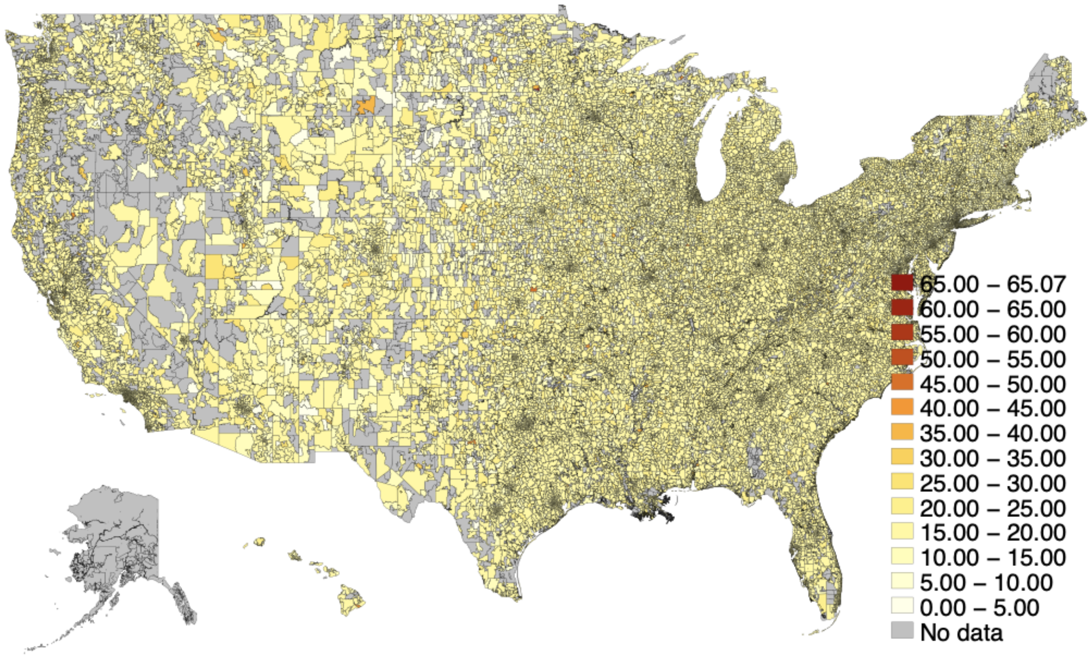
Through proximity to discontinuous points in lenders' credit functions, certain individuals face relatively higher uncertainty in obtaining a loan, ceteris paribus. We can disentangle two channels that guide our interpretation of the relationship between credit and voting. When the share of individuals below the salient thresholds of the lenders' credit function has a different effect on voting than the share above (polarization), we can infer that individuals are reacting to being credit constrained. In this case, the true relationship that we uncover is one between the tightness of credit constraints and the demand for higher deregulation. When the share of individuals below and above the salient thresholds have a similar effect on voting, we can instead presume that it is uncertainty in access to credit that relates to demand for deregulation. Our results – presented in section 3.5 – strongly support the latter.

Our main assumption is that, while the individual is not aware of the exact threshold used by her bank to approve her credit application, she can infer the proximity to the threshold by observing the successes and failures of her peers and her neighbors in obtaining credit. She thus perceives a relatively higher uncertainty in her ability to access credit when close to the salient threshold. We rely on the idea that social networks display homophily in characteristics used for computing credit scores and/or spatial correlation in credit scores. In our example, by noting that her social contacts or neighbors had somewhat different success on similar applications to hers (similar credit characteristics and similar homes), she infers that her application will face a substantial degree of uncertainty. This uncertainty would make her more likely to vote for those candidates who are more in favor of laxer requirements and regulations on the mortgage market front in order to increase her chance of obtaining credit. This is a direct economic effect of credit uncertainty on voting behavior. In Figure 3.1, we show the standard deviation of (log) credit score at the zip code level in 2010. It appears quite evident that the credit score varies little within zip code (the 50th and 99th percentile are 17% and 30% respectively). We find this consistent with our hypothesized *learning-about-uncertain-access* mechanism.[5]

In view of the recent literature on the political economy of voting (see section 3.1) we also conjecture that credit-access uncertainty – through its effect on home ownership, self-employment, and social mobility – might also fuel the fear of status loss. This non-monetary component of individual preference is central to the rhetoric of more conservative candidates in US politics (Guriev and Papaioannou (2022)) as the fear of status loss can be imputed to minority groups or foreigners following a "us against them" rhetoric. It has been shown that economic uncertainty and financial crises favor such "conservative shift" (Funke et al. (2016), Guiso et al. (2017)) and it is likely that the effect is salient for credit-uncertainty given the symbolic role of access to home ownership. Consequently, this "status effect" might also push voters facing credit uncertainty to favor more conservative candidates (democrats and republicans), irrespective of their economic positions

[5]While variations of at most 30% are quite limited, one can benchmark that versus the std. dev. of (log) income in the US in 2010 (25-65 years old) from the GRID database project which is 96% (`https://www.grid-database.org/`.

Figure 3.1: SD of (log) CS at ZIPCODE level (2010)



Notes: Legend is percentage variation within zip code in equally spaced 5pp intervals.

on redistribution (as in Rodrik and Mukand (2018)).

Consequently, we expect that the share of marginal-credit voters will affect the election results: the larger the share of credit-uncertain individuals the larger the voting shares for Republican candidates. We also expect that a larger share of credit-uncertain individuals will translate into more vote shares for socially conservative candidates, republicans and/or democrats.

3.3 Data

Our analysis combines several data sources: detailed individual-level data on credit reports from 2004 to 2016 from Experian (section 3.3.1); information on electoral outcomes for U.S. Congressional races from Leip (2017) and ideological positions of candidates in the liberal-conservative spectrum (Poole and Rosenthal, 1985, 1991) discussed in section 3.3.2; and several ancillary datasets such as trade exposure in local labor markets (Autor et al., 2020) and Census data (section 3.3.3).

3.3.1 Experian Credit Reports

We have access to proprietary data, provided by Experian, on a random 1% sample of the US population with valid credit scores in 2010, and yearly reports (drawn on June 30th of each year) for the same population from 2004 to 2016.

Similar data are described in detail in Lee and van der Klaauw (2010); and used in Albanesi and Nosal (2018) for the analysis of the effects of bankruptcy, Albanesi and Vamossy (2019) and De Giorgi et al. (2021) for the prediction of default and death, Herkenhoff et al. (2021) to analyze the impact of consumer credit on self-employment and entrepreneurship, in Mian and Sufi (2011), Adelino et al. (2016), Foote et al. (2021), and Albanesi et al. (2022) to analyze housing default of 2008. Bach et al. (2023) use the same data to study the lifecycle dynamics of credit. The credit report data contain all credit operations of individuals in the formal credit market and include credit scores, number and balances of revolving trades, mortgages, auto loans, credit limits on the different lines of credit, etc. Further, the data contain delinquencies and default events on each type of trade (it is common in the industry to refer to credit lines as trades, we will use the two interchangeably). Overall we have over 400 variables describing individual credit behavior for the entire period. In addition, the data contain some basic demographic information, i.e. date of birth, and zip code of residence.

These data contain two crucial pieces of information used in the construction of our main explanatory variable, detailed in section 3.4.2: credit scores and total credit amounts. Credit scores are computed using the Vantage Score V3 scoring model and are supposed to predict the probability of default in the next two years and rank individuals accordingly in a decreasing fashion between 300 and 850 score points. Figure 3.2 displays the distribution of credit scores in our data in different categories according to the scoring model. Our focus is on the population of borrowers with poor/fair credit scores as they are more likely to be marginal in the ability to access credit (section 3.4.2 elaborates on this choice). The total credit amount on open trades is the total amount of credit on all open accounts that an individual can potentially access (such as revolving credit, mortgages, and other loans). Figure 3.3 presents the average credit limits for consumers with different credit score bins, while 3.1 provides some additional statistics for the same bins. Limits differ substantially on average depending on the credit score bin, increasing with credit scores for all groups except for "excellent" borrowers (who are typically less leveraged and older). On average, those with very poor scores have a total limit of about 89,000USD, Poor/Fair about 120,000USD, Good 190,000USD, and Excellent 160,000USD. While credit scores are continuously distributed in the population, banks discontinuously increase total credit at locally determined thresholds along the credit score distribution (see Agarwal et al. (2017) and De Giorgi et al. (2023) for such examples). We focus on such thresholds at the commuting zone and year level as that is the lowest level at which credit markets might vary. Alternatively, we could have identified those thresholds at the State-year level – the typical level of banking regulation – however, it is not the appropriate level for the current paper as political candidates run for sub-State seats.[6]

Figure 3.2: Shares of individuals across the distribution of credit scores

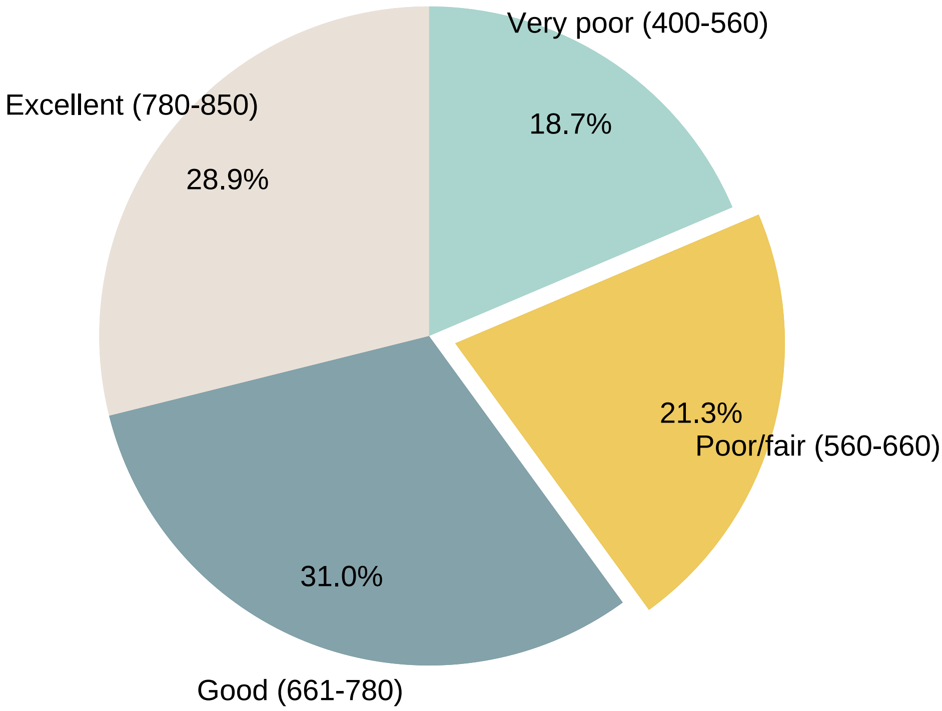


Table 3.1: Average Credit Limits in USD across the distribution of credit scores in the baseline year (2010)

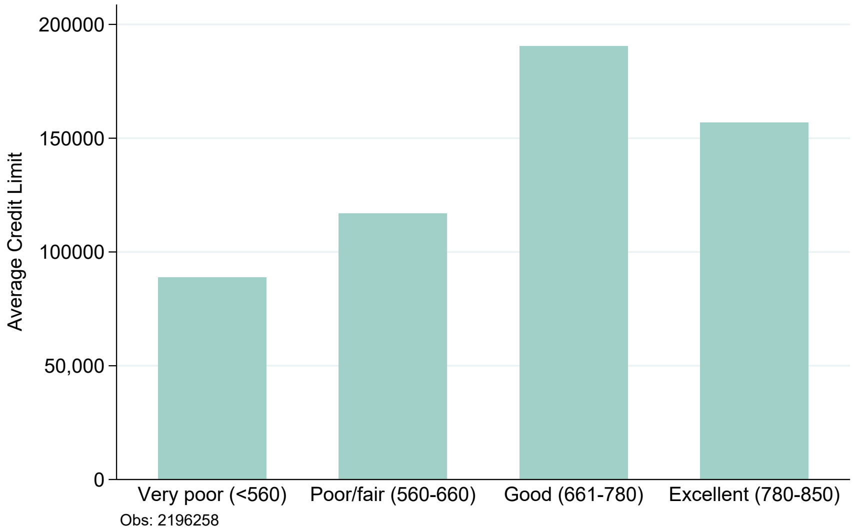
Group (credit score)	Mean	St. Dev.	Min.	Max.	Obs.
Very poor (<560)	88733.98	182381.37	0	15428795	230974
Poor/fair (560-660)	116946.38	201364.94	0	16246403	388954
Good (661-780)	190463.22	277642.29	0	17505724	619564
Excellent (780-850)	156859.70	207992.06	0	16773652	647251

3.3.2 Election Data

Data for electoral outcomes at the county-by-congressional district level between 2004-2016 are from Dave Leip's Atlas of U.S. Presidential Elections (Leip, 2017). These data track electoral outcomes over time for Presidential, House, and Senate elections in counties within electoral districts. We have the number of

[6] An alternative is to explore the probability of successful inquiries for new credit lines, as well as the size of mortgages. In the current paper, we focus on total credit amounts as it is the more complete series, and has the advantage of being determined for the large part on the simple credit score, while for mortgages only the typical approval process uses extensive information on the applicant which are not available to us.

Figure 3.3: Average Credit Limits across the distribution of credit scores in the baseline year (2010)



votes obtained by Democrats, Republicans, and other candidates in each electoral year. To maximize the frequency of the data, our main specification focuses on House of Representative elections. We choose these electoral data because the relevant geographic aggregation – county-by-congressional districts – allows us to incorporate information on local labor markets in our specification. We also use Poole-Rosenthal DW-NOMINATE scores to follow shifts in the ideological scores of candidates through roll-call votes in Congress (Poole and Rosenthal, 1985, 1991; McCarty et al., 2016). These scores represent legislator ideologies on a spatial map that allows for comparison across congresses. While several dimensions of these scores are available, we focus on the first dimension which intuitively represents the "liberal" vs. "conservative" divide of American politics. The scores vary from -1 (most liberal) to +1 (most conservative), with values at 0 representing the political center. These measures are widely used in political science and in particular for the study of polarization in the U.S. Congress (Persson and Tabellini, 2002; Bonica et al., 2013; Bonica, 2014; Fariss, 2014; Matsusaka, 1995; Autor et al., 2020).[7] To protect voter secrecy, electoral choices are not available at the individual level. Indeed, electoral outcomes at the precinct level are available (for example, Ansolabehere et al. (2014)), but they do not cover the timeline of our analysis and we were unable to attribute precinct vote shares to zip codes (the finest geographical information in our other data sources) and congressional districts. Other scholars rely on individually elicited voting preferences and behavior, through surveys or registration data. While this information is trackable at the individual level, it usually refers to small samples (N<4000) that are not repre-

[7] For greater detail see: `https://voteview.com/about`, `https://legacy.voteview.com/page2a.htm`.

sentative at the county-by-congressional district level or below (Enke, 2020; Mutz, 2018).

3.3.3 Other Data

This project leverages a wide array of ancillary datasets that are used for validation and support. To relate the Experian data – geolocalized at the 5-digit Zip Code Tabulation Area (ZCTA) level – with the electoral outcomes at the county-by-congressional district level, we build a crosswalk using 2010 ZCTAs to county FIPS relationship files, and ZCTA to Congressional District relationship files, both from the Census Bureau.[8] Throughout the paper, time-varying geographic units are restored to 2010 census tracks when possible as it coincides with the relevant census track for the sampling of the Experian data. Several ZCTAs are split across multiple counties. In order to ensure one-to-one mapping with counties, we attribute the zip code to the county where most of its population resides. This approach ensures little distortion as the population is usually concentrated in one county with minor spans over county borders (figure C1). We also ensure one-to-one mapping between ZCTAs and congressional districts by looking at official congressional directories for each Congress. We can then easily navigate between counties and congressional districts which do not coincide[9]

We obtain industry exposure to international trade from UN Comtrade, elaborated by David Dorn and described in Autor et al. (2020) [10] to account for shifts in voting behavior driven by local labor market exposure to international trade in wake of China's entry into the WTO in 2001. These data contain both local labor market exposure to trade with China, as well as trade flows between China and other high-income countries with trade flows comparable to the US.[11] These data will be used in the main specification following Autor et al. (2020) to account for the causal effect of increased exposure to international trade on political polarization. Population data from the US Census Bureau complete the data we use to include information on the racial, gender, and age composition of our geographic units of interest.

A standard concern in studies of political outcomes is gerrymandering, the strategic redistricting of incumbent policymakers after each census to favor the incumbent party in future elections. Following Autor et al. (2020), we use county-by-congressional districts as our main geographic unit, as congressional districts

[8] `https://www.huduser.gov/portal/datasets/usps_crosswalk.html`

[9] `https://www.govinfo.gov/app/collection/cdir`

[10] `https://www.ddorn.net/data.htm#Industry%20Trade%20Exposure`, sections D, E.

[11] The eight other high-income countries have trade data comparable to the US over the full sample period: Australia, Denmark, Finland, Germany, Japan, New Zealand, Spain, and Switzerland.

are designed along population criteria and often span across counties and commuting zones. There are ultimately 435 congressional districts that are designed to contain the same number of voters. Section C.2 discusses how we break down the data to ensure that the geographic limits of our county-by-congressional district units remain constant.

3.4 Empirical design

This section presents our main specification and our identification strategy. It also details the construction of our main explanatory variable measuring credit uncertainty and provides insight into the identification strategy.

3.4.1 Main specification

Our aim is to investigate the relationship between uncertainty in credit access and voting behavior. We conjecture that voters' uncertainty about their ability to access credit, represented by their proximity to "random" thresholds that we detail in subsection 3.4.2, pushes voters toward candidates that push for laxer credit market regulation and more populist ones.

In order to establish such a link, we estimate the following equations:

$$\begin{aligned}
\texttt{Vote Share}_{c,t} &= \alpha_1 + \beta_1 \texttt{Share at Thresholds}_{c,t} + \gamma_1 \mathbf{X}_{c,t} + \mathbf{D}_c + \mathbf{D}_t + \epsilon_{1,c,t} && (3.1)\\
\texttt{Vote Share}_{c,t} &= \alpha_2 + \beta_2^b \texttt{Share Below Thresholds}_{c,t} \\
&\quad + \beta_2^a \texttt{Share Above Thresholds}_{c,t} + \gamma_2 \mathbf{X}_{c,t} + \mathbf{D}_c + \mathbf{D}_t + \epsilon_{2,c,t} && (3.2)
\end{aligned}$$

where the left-hand-side variable, $\texttt{Vote Share}_{c,t}$, is the Republican (or Democrat) vote share in county-by-congressional district c and election year t. Our analysis covers the US county-by-congressional house elections over the 2004-2016 period. `Share at Thresholds`, `Share below Thresholds`, and `Share above Thresholds` denote the share of voters around, below, and above the most salient credit score thresholds in district c and year t (described in section 3.4.2 and Table 3.3). $\mathbf{X}_c$ denotes a set of time-varying district-specific covariates, such as the gender and race composition of the district and the (instrumented) share of China imports in the district (described in section 3.3). $\mathbf{D}_c$ and $\mathbf{D}_t$ denotes the full set of county-by-congressional district and time fixed effects, respectively. The choice of this specification is made necessary because we do not observe individual voting behavior and candidates vary within congressional districts, while credit markets likely vary at most at the local labor market level (i.e. groups of adjacent counties). By aggregating vote counts, personal finance information, and demographic data at the county-by-congressional district level, we are effectively estimating equations (3.1) and (3.2) at the finest possible level.

The βs are our main coefficients of interest, representing the causal relationship between increases in the share of individuals who are uncertain in their ability to borrow and electoral outcomes. We exploit time and geographic variation in these shares driven by exogenous supply-side shifts in lenders' credit functions to identify such effects. Section 3.4.2 describes how we identify and validate the discontinuities in lenders' credit functions, while appendix C.1 complements the framework with extensive sensitivity analyses of these methods.

An alternative specification uses DW-NOMINATE scores measuring candidate ideological positions along the traditional progressive-conservative spectrum as outcomes of interest. Except for the left-hand-side variable, everything else remains unchanged. The equations with the ideology measures as outcomes are additionally estimated separately for democrat- and republican-winning districts to detect polarization, i.e. shifts away from the political center. These specifications complement the vote share equations as they elicit the effect of uncertainty in access to credit on the intensity of ideological positions beyond bipartisanship.

3.4.2 Credit Thresholds

A crucial step of our analysis is to find plausible sources of exogenous variation in credit access to then build our main explanatory variable: the share of individuals close to the threshold. We interpret proximity to a salient threshold as generating high uncertainty in credit access. To identify the thresholds for credit access we employ an approach similar to Agarwal et al. (2017).

Variation in access to credit along the values of credit scores is plausibly exogenous as the density of credit scores is smooth at the thresholds (Figure 3.5 and Figure C3) and credit scores are locally volatile, such that individuals cannot exactly manipulate their score. The exact formula for the computation of the credit score is proprietary to Vantage. While it is known what type of credit information goes into that calculation, consumers cannot exactly point towards a specific score.[12] In addition, it is very common to have one's credit score move up or down by 5 points even within a month despite little changes in fundamental behavior, for example, that would happen if there is an increase or fall in credit card balances (even by small amounts), opening or closing credit lines, etc.

The thresholds are identified with a simple regression discontinuity setup at

[12] https://vantagescore.com/press_releases/the-complete-guide-to-your-vantage score/

the commuting zone level in each election year.[13] We test for discontinuities in the total credit limit at five-point intervals in credit scores and we define the thresholds that determine discontinuous access to credit as the credit scores for which we detect an increase in the credit limit in each local regression. Eventually, we compute the explanatory variable used in the main specification as the share of individuals close to the threshold with respect to the total population in each commuting zone and election year.

We run the regressions in each commuting zone CZ and election year t. Let c_i be the credit score of individual i, y_i be the total credit limit, d be a dummy equal to 1 if it is above the cutoff value of the credit score $\bar{c}$ and 0 otherwise, and $\mathbf{D}_c$ denote county fixed effects. Following Cattaneo et al. (2016), we estimate

$$\sinh^{-1}(y_i) = \begin{cases} \alpha d_i + \beta_{01}(c_i - \bar{c}) + \beta_{02}(c_i - \bar{c})^2 + ... + \beta_{0p}(c_i - \bar{c})^p + \mathbf{D}_c & \text{if } c_i < \bar{c} \\ \alpha d_i + \beta_{11}(c_i - \bar{c}) + \beta_{12}(c_i - \bar{c})^2 + ... + \beta_{1p}(c_i - \bar{c})^p + \mathbf{D}_c & \text{if } c_i \geq \bar{c} \end{cases} \tag{3.3}$$

for all $\bar{c}$ in 5-point intervals between 560 and 660, election year, and commuting zone.[14] The degree of the polynomial transformation p is optimally chosen to allow for the potential outcome to have some direct dependence on the credit score and is equal to 4 in our setting (Calonico et al., 2014; Cattaneo et al., 2020). We interpret positive and statistically significant $\alpha(\bar{c})$ as discontinuous and locally exogenous increases in access to credit where $\bar{c}^*$ is the threshold. Commuting zones represent local labor markets within which banks decide their credit functions competitively and are appropriate markets due to the large number of individuals who access banking services through credit unions which are local in nature.[15] Furthermore, as credit-constrained individuals are mostly located at the bottom of the distribution of credit scores, they are more likely to be serviced by CUs that guarantee minimum financial services (Office of Small Credit Union Initiatives, 2016). Although banks need not choose discontinuous credit functions, the data and the literature suggest this is the case (Agarwal et al., 2017; De Giorgi et al., 2023). To ensure sufficient power, we restrict our analysis to commuting zones with more than 500 observations. We further restrict the portion of the distribution of credit scores within which we search for thresholds in the 560-660 interval. We

[13]Our data would allow us to estimate the thresholds at a yearly frequency. As the outcome of interest in the main specification varies biannually (electoral results), we focus on identifying the thresholds at this level. To ensure sufficient power in sparsely populated commuting zones, we exploit observations from both election and non-election years for estimation.

[14]Cutoffs $\bar{c}$ vary at the commuting zone and yearly level ((CZ, t) in our notation). For simplicity, we drop these subscripts hereafter.

[15]Between 2004 and 2016, 94 million customers on average were members of credit unions. Even though the number of credit unions has been decreasing, there were 7944 CUs on average (National Credit Union Administration, 2016).

choose 660 as the upper limit of our interval of interest because it represents the cutoff between "fair" and "good" credit scores in the Vantage credit score model.[16] We do not seek thresholds above 660 because fewer individuals would be marginal at these values. We do not seek thresholds below 560 because the distribution of credit limits is distorted by the fact that individuals who face solvency issues are pushed down to credit scores around 525 independently of their previous credit score, which creates bunching and negative thresholds because of extraordinary behavior, such as declaring bankruptcy. 21.3% of individuals in our dataset exhibit credit scores in this range (as a reminder figure 3.2 provides additional information on the distribution of credit scores).

We successfully run 59,237 regressions, i.e. in election year-commuting zone units with more than 500 observations. For each election year, we have between 395 and 410 commuting zones and 298-316 valid thresholds $\bar{c}^*$, i.e. cutoffs for which $\alpha_{\bar{c}^*}$ is positive and statistically significant. In certain instances, we identify positive and statistically significant thresholds in certain years and not in others. When a commuting zone presents discontinuities in access to credit in one year, we impute the missing election years using previous observations under the assumption that banks' decisions to grant credit according to discontinuous credit functions do not vary between election years. In a few instances, we are never able to detect discontinuities over time for certain commuting zones and thus disregard them. Table 3.2 summarizes these results and the available thresholds. We also detect multiple or contiguous thresholds in certain commuting zones and election years. In these cases, we only keep the threshold with the largest coefficient, thus implying the largest jump in credit limit. We identify thresholds according to this conservative rule to reduce the chance of detecting "false thresholds" due to discontinuity detection bias, for example, due to variation in the density of borrowers around the thresholds and discontinuities in credit demand.

Figure 3.4 provides additional information on the location of the 2,821 thresholds $\bar{c}^*$. While thresholds around 560 and 630 are more frequently observed in the dataset, all values of $\bar{c}^*$ are present.

Figure 3.5 displays the (average) magnitude of the discontinuities in the total credit limit that we detect around the thresholds using equation (3.3). For each election year, we center observations around the relevant threshold in the individual's commuting zone and plot the density of individuals in each credit score point (green bars, left axis). While the density slightly increases with higher credit scores, there is no bunching of individuals just above or below the thresholds, so the density is smooth at the threshold (McCrary (2008) test). The black lines re-

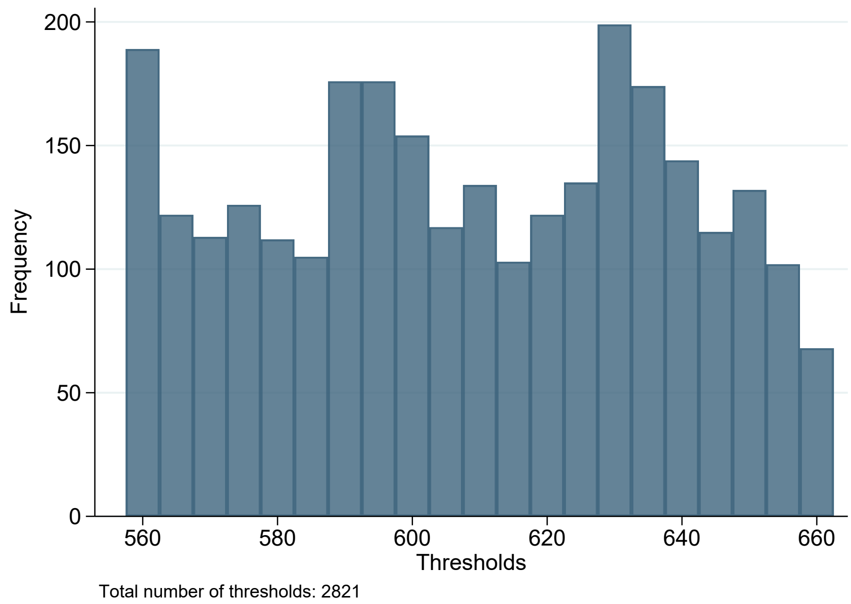
[16] `https://www.experian.com/blogs/ask-experian/credit-education/score-basics/what-is-a-good-credit-score/`

Table 3.2: Commuting Zones (CZ) and valid thresholds

Election Year	Total CZ	CZ with thresh.	CZ with thresh. (imp.)
	(1)	(2)	(3)
2004	409	302	408
2006	410	303	408
2008	408	312	408
2010	403	309	403
2012	396	311	396
2014	400	316	400
2016	395	298	395

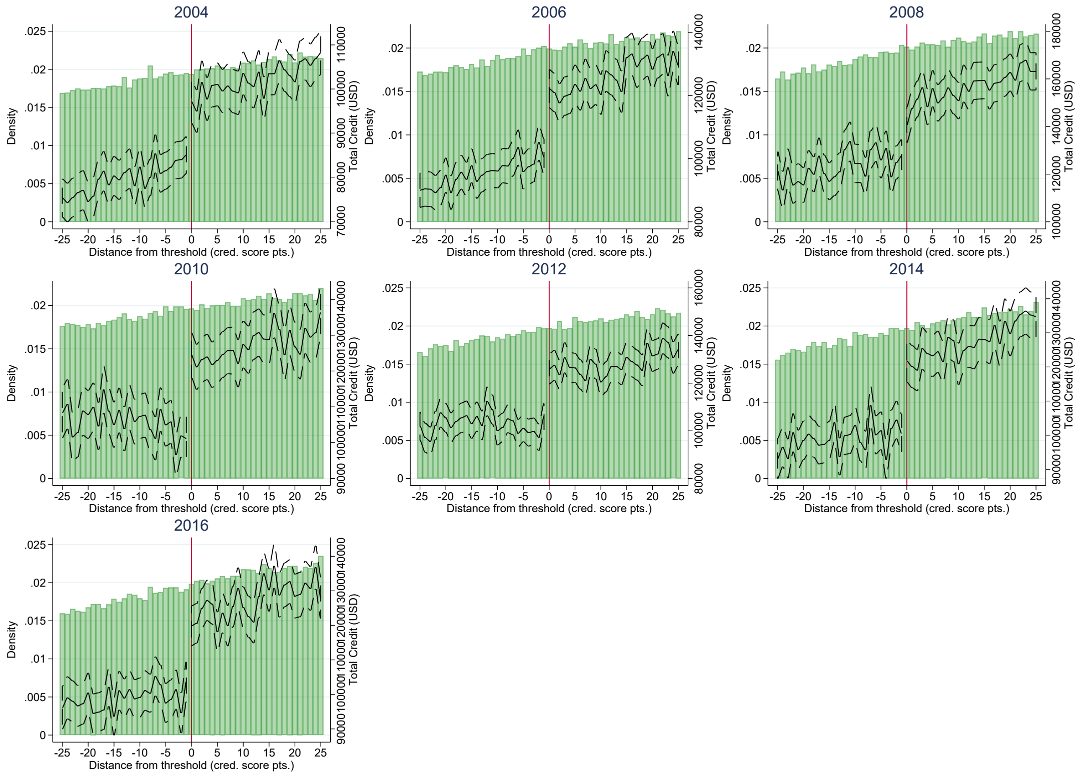
(1) – total CZs with sufficient observations (i.e. > 500), (2) – number of CZs for which we detect a positive and statistically significant $\bar{c}^*$, (3) – number of CZs for which we are able to impute missing thresholds by keeping previous ones in years in which no valid threshold was detected.

Figure 3.4: Frequency of thresholds $\bar{c}^*$



flect our estimates of credit limits from equation (3.3) with 95% CIs, transformed back into dollar values (right axis). In all election years, the discontinuity in credit limits around the thresholds is statistically significant and large, increasing on average by 20,000USD around the thresholds. On average, valid thresholds have a coefficient $\alpha = 1.14$ (inverse hyperbolic sine scale). This implies an increase in the credit limit of 23,400USD as low as the 25th percentile of the distribution of borrowers (from 11,013USD to 34,435USD), and larger increases thereafter. As credit limits are automatically updated by lenders and the jumps are economically large, it is unlikely that discontinuity detection biases alone explain these jumps. Figures C2 and C3 in the appendix provide additional evidence on the discontinuity of credit limits (Figure C2) and of the smoothness of credit scores (Figure C3) around the centered thresholds.

Figure 3.5: Thresholds: density of individuals and total credit amount



3.4.2.0.1 The main explanatory variable. Once we have identified the thresholds $\bar{c}^*$, we can build the main explanatory variable as the share of individuals with a credit score close to the threshold with respect to the total population in the county by congressional district cell as used in equations (3.1) and (3.2). No crosswalks are needed as commuting zones are groups of contiguous counties that form local labor markets and thus fully contain our geography of interest, county-by-congressional districts (CCD). The main assumption that underlines our identification of the discontinuities in credit limits is that being above or below the threshold is locally random. In fact, it is actually weaker than that, we simply need that the share of people around those thresholds are quasi-random. That appears rather plausible as consumers do not know where the thresholds are.

As we have no specific size for selecting the neighborhood of the thresholds in terms of credit score points, we compute the shares of individuals close to the thresholds within 5, 10, 15, 20, and 25 credit point deviations above, below, and in total. In our preferred specification, we will analyze shares with a bandwidth of 15 credit score points. Table 3.3 summarized the population shares of individuals above, below, or close to the thresholds. Not surprisingly, the shares increase as the bandwidth of credit score points considered increases. As expected, roughly the same amount of individuals are above and below the thresholds, further supporting that there is no bunching. In the preferred specification, 6.6% (std.dev. 4.2%) of individuals are within 15 credit score points of the threshold.

Table 3.3: Population shares close to 500s thresholds at different bandwidths (BW) and weighted by population (14,549,479 total observations)

Variable	Mean	St. Dev.	Min	Max
share(tot), BW: 5	0.023	0.026	0	1
share(above), BW: 5	0.013	0.020	0	1
share(below), BW: 5	0.011	0.016	0	1
share(tot), BW: 10	0.045	0.035	0	1
share(above), BW: 10	0.024	0.026	0	1
share(below), BW: 10	0.021	0.022	0	1
share(tot), BW: 15	0.066	0.042	0	1
share(above), BW: 15	0.035	0.030	0	1
share(below), BW: 15	0.032	0.028	0	1
share(tot), BW: 20	0.088	0.048	0	1
share(above), BW: 20	0.046	0.034	0	1
share(below), BW: 20	0.042	0.032	0	1
share(tot), BW: 25	0.110	0.054	0	1
share(above), BW: 25	0.058	0.037	0	1
share(below), BW: 25	0.052	0.036	0	1

3.4.3 Discussion on identification.

Section 3.4.2 clarifies how we identify the cutoff points, the points in the credit score distribution – for poor and fair borrowers – where a few score points make a substantial difference in the probability of getting a loan and its size. Over the whole sample, we estimate an average 20,000USD jump in the total credit available to individuals on either side of the cutoffs. These jumps are substantial, typically larger than 20% of the total credit limit. Customers/voters do not know where these thresholds are exactly, they are typically not communicated and depend on the internal model of banking used by each bank at the local level. Banks also experiment with such thresholds in order to acquire new customers (De Giorgi et al. (2023)). What is relevant for our analysis is that voters in those proximities are aware that obtaining a loan is not a done deal. In fact, the outcome of a mortgage or credit application is rather uncertain as it will depend on where exactly the thresholds are. Exposed customers face substantial uncertainty in their ability to borrow. We hypothesize that while customers do not exactly know where the cutoffs are, they are aware that they are in a region of the credit score distribution where obtaining additional credit is in doubt, and therefore operate under larger uncertainty than consumers with larger, inframarginal, credit scores. Importantly, the fact that the discontinuities are not known by the individuals does not pose a threat to the identification of the thresholds, as the lack of common knowledge about the thresholds' location does not affect the efficiency of the treatment effect estimator asymptotically (Porter and Yu, 2015).

We remind the reader that shifts of a few score points within a short period of time are rather frequent, such that the monthly individual variation for those between 560 and 660 credit score points is typically plus/minus 5 points without much action on the credit side. While customers have some control over their scores, they cannot pinpoint them. In addition, they have no knowledge of where exactly the thresholds are and when exactly their credit score will change, as it depends on when the relevant financial institutions send in their reports to the credit bureau. Furthermore, we note that credit score apps, and costless credit score verification, were not overly present before 2010, and checking personal scores would have been costly as a small fee would be charged. As itemized credit reports were not widely available, actions to improve credit scores were also harder to single out.[17]. Consequently, credit-constrained individuals who face uncertainty in their ability to borrow would inquire frequently with the lenders and thus observe their own success rate and that of their peers (although imperfectly). Because of the frequent variation of the credit score with respect to the threshold and the infrequent nature of elections, we believe that the relative position of the individual with respect to the threshold – as observed in June of the election year – is effectively random.

Finally, the βs in equations (3.1) and (3.2) are our coefficients of interest. They inform us of how uncertainty in access to credit drives shifts in the vote margin and the ideological position of winning candidates through variation across time and space in the share of individuals around the thresholds. We allow for the effect of uncertainty on polarization to be heterogeneous depending on the share of individuals below and above the threshold, separately (equation (3.2)). The separation of specifications (3.1) and (3.2) guides the interpretation of the coefficients. Heterogeneous behavior above and below the threshold (β_2^b and β_2^a) is consistent with individuals knowing their position relative to the threshold, therefore eliciting the effect of exogenous variation in credit constraints on voting. Homogeneous behavior is consistent instead with uncertainty in credit access: the small fluctuations in

[17]For a brief history of the leading app in this market, Credit Karma, see `https://www.creditkarma.com/ourstory`

credit scores that affect the individual's borrowing capacity cannot be controlled and individuals notice that similar borrowers obtain higher/lower credit limits and cannot control their status relative to them. Our results showcase homogeneous above- and below-threshold behavior and are thus consistent with the uncertainty in access to credit narrative. These findings are discussed in-depth in section 3.5.

3.5 Voting

3.5.1 Main results - Voting Shares

Table 3.4 displays our baseline results produced by the estimation of equation (3.1). Population weights with the counts of individuals in each cell used to compute the shares of individuals close to the thresholds are included and all standard errors are clustered at the county-by-congressional district level. Column 1 displays the effect on Republican vote share (in percentage) of a variation in the share of people just above (line 2) and below the credit threshold (line 3) by a margin of 15 credit points. Column 2 displays the effects on republican voting of an increase in the share of people within a margin of 15 points above and below the credit threshold. Both specifications show a strong and significant positive effect. In county-by-congressional district cells with a higher share of people clustered around the salient credit threshold (with higher credit uncertainty), we observe a higher share of republican votes than in other areas. In particular, a 10 percentage point increase in the share close to the thresholds will increase republican vote share by 2.7pp and decrease democratic votes by 2.6pp. These findings are consistent with the hypothesis that proximity to salient thresholds increases uncertainty in access to credit, as opposed to differentiating between individuals who face varying credit constraints. All columns control for the (instrumented) share of China imports following the specification of local trade exposure in commuting zones derived by Autor et al. (2014, 2020); Acemoglu et al. (2016), as well as the percentage of white people as well as the share of women in the district.[18] The coefficients for these demographic controls are of the expected signs: a higher share of white and male voters in an area translates into a higher vote share for republicans. The effect is the opposite for democrats, benefiting more from a higher share of non-white and women voters. Interestingly, the main effect of proximity

[18]We account for shocks in local labor markets driven by the growth of Chinese import penetration by measuring trade flows between the U.S. and China at the local labor market level, accounting for industry structure within each LLM at the beginning of the analysis. To disentangle demand and supply shocks to LLMs, we instrument the import exposure variable with the composition and growth of Chinese imports in eight other developed countries with comparable trade data: Australia, Denmark, Finland, Germany, Japan, New Zealand, Spain, and Switzerland, following Autor et al. (2020).

to the credit threshold in column 2 is very similar above and below the threshold in column 1: what seems to matter for voting behavior is the overall proximity to the threshold – and the uncertainty it generates – more than the direct economic effect created by the jump above the threshold.

Columns 3 and 4 display similar specifications for democratic voting. The effects are the close opposite of what we observed in columns 1 and 2, which is not surprising given the almost perfect two-party system save for independent candidates: an increase in the share of people below or above the threshold (column 3) and around the threshold (column 4) is associated with a decrease in democrat vote shares in these areas. As for column 2, we notice that the effect is similar for the share of people below and above the threshold: the effect is driven by the overall proximity to the threshold and not the crossing of the threshold itself. The coefficient displayed in column 4 implies that an increase of 10pp in the share of people with credit scores located around the threshold implies a decrease in democrat vote shares of 2.6pp.

Table 3.4: Vote Shares and Credit Access Uncertainty

VARIABLES	(1) Rep	(2) Rep	(3) Dem	(4) Dem
Share close thresh.		0.268**		-0.264**
		(0.119)		(0.118)
Share close thresh. Above	0.248**		-0.258**	
	(0.126)		(0.125)	
Share close thresh. Below	0.292*		-0.272*	
	(0.158)		(0.156)	
Share China Import	-0.019	-0.019	0.017	0.017
	(0.014)	(0.014)	(0.014)	(0.014)
Share White	0.705***	0.705***	-0.698***	-0.698***
	(0.020)	(0.020)	(0.020)	(0.020)
Share Female (voting age)	-0.969***	-0.969***	0.967***	0.967***
	(0.183)	(0.183)	(0.179)	(0.179)
Fixed effects	County x congressional district, year			
Observations	12,053,931	12,053,931	12,053,931	12,053,931
R-squared	0.435	0.435	0.420	0.420
Mean of dep. var.	0.464	0.464	0.505	0.505

*** p<0.01, ** p<0.05, * p<0.1

Cluster robust SE at county-by-congressional district in parenthesis.

3.5.2 Conservative shift of elected candidates.

Table 3.5 displays results where we look at the effect of the share of potential voters around the credit threshold on the political position of the elected candidate (using the well-known DW-NOMINATE index developed by Keith T. Poole and Howard Rosenthal (Poole and Rosenthal, 1985, 1991; McCarty et al., 2016)). The main objective is to test the "ideological channel" purporting that credit uncertainty pushes voters toward candidates using a more conservative rhetoric through, for example, a fear of status loss. Recall that the main dimension of DW-NOMINATE scores locates candidates on a "liberal-conservative" space, ranging from -1 (most liberal) to +1 (most conservative) (see section 3.3). To perform this analysis we use the specification in Equations 3.1 where we replace the left-hand-side variable with the DW-NOMINATE index. Columns 1 and 2 present the results for the Republican-winning congressional districts splitting coefficient for the share of citizens above and below the threshold (column 1) and around the threshold (column 2). We observe that the share of people affected by credit uncertainty does not impact the ideology of elected candidates in republican-winning districts, while trade exposure to the "China shock" seems to matter significantly (as in Autor et al. (2020)). However, columns 3 and 4 show that for Democratic-winning congressional districts, our measure of credit uncertainty explains an increase toward more conservative ideology by a substantial margin. This effect on the ideology of elected candidates is again symmetric with respect to the share of people above and below the threshold. Finally, columns 5 and 6 confirm that the conservative shift of elected candidates triggered by the share of people facing credit uncertainty is true for all congressional districts, irrespective of party affiliation. Column 6 implies that an increase of 10pp of the share of people located around a salient credit threshold increases by 5pp the DW-NOMINATE score of the average elected candidate. This is a substantial increase as it represents a doubling of the mean of the index (equal to 0.051 for all candidates) on the direction of more conservative ideology.

This set of results paints an overall consistent picture: an increase in the share of people facing credit uncertainty is associated with a shift toward the election of more conservative candidates, and this effect is particularly true for Democratic candidates. This substantiates the idea that economic uncertainty fuels conservative rhetoric by reinforcing the "within-group" narratives and identifying external groups as the main causes of economic hardship (e.g. by pointing to the fact that a specific minority group benefits more from a specific housing program). We read this result as consistent with the literature discussed in the Introduction (Bonomi et al. (2021),Grossman and Helpman (2020)) showing that economic shocks have the potential to magnify the political salience of racial and ethnic identity, yielding an increase in conservative voting on social issues even conditional on economic status.

Table 3.5: DW-NOMINATE scores along the progressive (-) – conservative (+) dimension and exposure to credit uncertainty

VARIABLES	(1) Rep.	(2) Rep.	(3) Dem.	(4) Dem.	(5) All	(6) All
Share close thresh.		-0.035		0.471***		0.509**
		(0.080)		(0.145)		(0.258)
Share close thresh. Above	-0.134		0.478***		0.434	
	(0.082)		(0.168)		(0.269)	
Share close thresh. Below	0.104		0.465**		0.599*	
	(0.117)		(0.186)		(0.344)	
Share China Import	0.023***	0.023***	0.014	0.014	0.030	0.030
	(0.008)	(0.008)	(0.018)	(0.018)	(0.029)	(0.029)
Share White	0.048*	0.047*	0.286***	0.286***	1.103***	1.103***
	(0.025)	(0.025)	(0.021)	(0.021)	(0.048)	(0.048)
Share Female (voting age)	0.146	0.146	-0.409*	-0.409*	-1.037**	-1.037**
	(0.184)	(0.184)	(0.237)	(0.237)	(0.434)	(0.434)
Fixed effects	County x congressional district, year					
Observations	6,275,158	6,275,158	5,867,112	5,867,112	12,146,679	12,146,679
R-squared	0.053	0.052	0.228	0.228	0.296	0.296
Mean of dep. var.	0.454	0.454	-0.380	-0.380	0.051	0.051

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

Clustered SE at county x congressional district, 15 credit score point bandwidth.

(1)-(2) ideological measures for Republican elected representatives, (3)-(4) Democrat elected representatives, (5)-(6) all elected representatives.

3.5.3 Sensitivity analysis

We perform a series of robustness exercises on our baseline results in C2-C10. Tables C2-C3 display results for the share of people above and below the threshold where we vary the bandwidths around the thresholds from 5 to 25 credits points by increments of 5 points (recall that our baseline results are based on a bandwidth of 15 points). Table C2 displays the results for democratic vote shares and Table C3 for Republican vote shares. In both tables, the main coefficients remain very stable as we vary the size of the bandwidth. Tables C4-C5 provide similar robustness checks for the share of people around (above and below) the threshold, for republican and democratic vote shares. We observe in these tables a decrease in the magnitude of the effect as we increase the bandwidth. This is intuitive, as increasing the bandwidth dilutes the uncertainty associated with proximity to the credit threshold and most likely decreases the strength of our main effect.

Tables C6, C7, C8 propose similar robustness exercises, with varying bandwidth, focusing on the DW-NOMINATE score of the winning candidate. Tables C6 and C7 display results for various bandwidths above/below the thresholds and around the thresholds, for republican and democratic-winning candidates. The results in Table C6 are consistent with our baseline results: increasing the share of individuals around salient credit thresholds tends to be associated with more conservative democrat-winning candidates (above and below). Table C7 shows that the share of people around the threshold (above and below) tends to render republican-winning candidates more conservative. Interestingly, table C8 shows that our main variable of interest has a strong impact on the conservative rhetoric of all elected candidates, republican and democrat. Similarly to the results on vote shares, the effect is smaller as the bandwidth increases.

We then turn toward to possible confounding role of gerrymandering, as some of the congressional districts were redesigned over our period of analysis. Table C.2 shows results only for 2012, 2014 and 2016 elections, where no redistricting took place. Our baseline results are qualitatively robust to this specification, indicating that gerrymandering per se does not drive the results of our main specification. Focusing on those election years for the DW-NOMINATE score (table C10), all columns show that the main coefficients tend to display a lower significance, while similar in magnitude.

3.6 Concluding Remarks

We show a novel role of credit access in determining political and voting behavior. Credit access has a strong economic and symbolic value, as it is an essential component of the American dream through the social ladder of housing, and self-employment. We highlight that individuals with uncertain access to credit – located around salient credit score thresholds – appear to disproportionately vote for republican candidates and more conservative candidates overall.

In particular, we show that a 10pp increase in the share of individuals at the margin of credit access causally increases the vote shares for Republicans by 2.7pp (and decreases democratic vote shares by 2.6pp) a margin which would in several instances change the election results. Further, we show that even in democratic Congressional districts an increase in the share of marginal credit individuals renders democratic candidates' rhetoric more conservative. An increase of 10pp in the share of individuals around the threshold doubles the average of the NOMINATE score across all candidates, indicating a strong move toward a more conservative ideology.

We interpret our results as consistent with 2 channels. First, a direct economic channel, as higher uncertainty in credit access might favor Republican candidates who typically run on platforms of deregulation and easier access to credit. Second, our results also support a complementary "cultural channel", as economic uncertainty can amplify the political salience of racial and ethnic identity and create a conservative shift (see Bonomi et al. (2021) or Grossman and Helpman (2020)). This is what we observe with the DW-NOMINATE score, especially for elected Democratic candidates.

These results bear important policy implications as credit access is usually easier to manipulate at the local level than other macroeconomic indicators (e.g. trade access, GDP, etc.). Our results also highlight that political conservatism might find part of its origin in credit uncertainty. Legislators can therefore ease access to credit (or reduce credit uncertainty) through local regulations and avoid an unwanted rise in social conservatism and racial tensions.

Future work can look at the exact functioning of the cultural channel. In particular, future research might study in more detail if the DW-NOMINATE score varies differentially with credit uncertainty in areas with different racial and gender compositions. It might investigate if the effect on conservatism is more pronounced in areas with stronger regulation that favors discriminated minorities (e.g. housing regulation).

Bibliography

Acemoglu, D., Autor, D., Dorn, D., Hanson, G. H., and Price, B. (2016). Import competition and the great us employment sag of the 2000s. *Journal of Labor Economics*, 34(S1):S141–S198.

Acemoglu, D., He, A., and le Maire, D. (2022). Eclipse of rent-sharing: The effects of managers' business education on wages and the labor share in the us and denmark. Technical report, National Bureau of Economic Research.

Adams, R. B. and Kirchmaier, T. (2016). Women on boards in finance and stem industries. *American Economic Review*, 106(5):277–81.

Adelino, M., Schoar, A., and Severino, F. (2016). Loan Originations and Defaults in the Mortgage Crisis: The Role of the Middle Class. *The Review of Financial Studies*, 29(7):1635–1670.

Agarwal, S., Chomsisengphet, S., Mahoney, N., and Stroebel, J. (2017). Do Banks Pass through Credit Expansions to Consumers Who want to Borrow?*. *The Quarterly Journal of Economics*, 133(1):129–190.

Ahn, T., Arcidiacono, P., Hopson, A., and Thomas, J. R. (2019). Equilibrium grade inflation with implications for female interest in stem majors. Technical report, National Bureau of Economic Research.

Aina, C. and Nicoletti, C. (2018). The intergenerational transmission of liberal professions. *Labour Economics*, 51:108 – 120.

Albanesi, S., DeGiorgi, G., and Nosal, J. (2022). Credit growth and the financial crisis: A new narrative. *Journal of Monetary Economics*, 132:118–139.

Albanesi, S. and Nosal, J. (2018). Insolvency after the 2005 bankruptcy reform. Technical report, National Bureau of Economic Research.

Albanesi, S. and Vamossy, D. F. (2019). Predicting consumer default: A deep learning approach. Technical report, National Bureau of Economic Research.

Algan, Y., Guriev, S., Papaioannou, E., and Passari, E. (2017). The European trust crisis and the rise of populism. *Brookings Papers on Economic Activity*, 48(2):309–400.

AlmaLaurea, C. I. (2020). Note metodologiche: Xxii indagine – profilo dei laureati 2019. Technical report.

AlmaLaurea, C. I. (2021a). Note metodologiche: Xxiii indagine – profilo dei laureati 2020. Technical report.

AlmaLaurea, C. I. (2021b). Xxiii indagine – profilo dei laureati 2020. Technical report.

Altonji, J. G. (1993). The demand for and return to education when education outcomes are uncertain. *Journal of Labor Economics*, 11(1, Part 1):48–83.

Altonji, J. G., Arcidiacono, P., and Maurel, A. (2016). The analysis of field choice in college and graduate school: Determinants and wage effects. In *Handbook of the Economics of Education*, volume 5, pages 305–396. Elsevier.

Altonji, J. G., Blom, E., and Meghir, C. (2012). Heterogeneity in human capital investments: High school curriculum, college major, and careers. *Annu. Rev. Econ.*, 4(1):185–223.

Altonji, J. G. and Zhong, L. (2021). The labor market returns to advanced degrees. *Journal of Labor Economics*, 39(2):303–360.

Altonji, J. G. and Zimmerman, S. D. (2018). The costs of and net returns to college major. In *Productivity in higher education*, pages 133–176. University of Chicago Press.

Amemiya, T. (1985). *Advanced econometrics*. Harvard university press.

Anderson, M., Brown, R., Kerwin, C., and Rees, D. (2020). Occupational licensing and maternal health: Evidence from early midwifery laws. *The Journal of Political Economy*, 128(11):4337–4383.

Angelucci, M., De Giorgi, G., Rangel, M. A., and Rasul, I. (2010). Family networks and school enrolment: Evidence from a randomized social experiment. *Journal of Public Economics*, 94(3-4):197–221.

Angrist, J. D. and Guryan, J. (2008). Does teacher testing raise teacher quality? Evidence from state certification requirements. *Economics of Education Review*, 27(5):483 – 503.

Ansolabehere, S., Palmer, M., and Lee, A. (2014). Precinct-level election data.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics*, 121(1-2):343–375.

Arcidiacono, P., Aucejo, E., Maurel, A., and Ransom, T. (2016a). College attrition and the dynamics of information revelation. Technical report, National Bureau of Economic Research.

Arcidiacono, P., Aucejo, E. M., and Hotz, V. J. (2016b). University differences in the graduation of minorities in stem fields: Evidence from california. *American Economic Review*, 106(3):525–62.

Arcidiacono, P., Cooley, J., and Hussey, A. (2008). The economic returns to an mba. *International Economic Review*, 49(3):873–899.

Arcidiacono, P., Ellickson, P. B., et al. (2011). Practical methods for estimation of dynamic discrete choice models. *Annual Reviews of Economics*, 3(1):363–394.

Ashworth, J., Hotz, V. J., Maurel, A., and Ransom, T. (2021). Changes across cohorts in wage returns to schooling and early work experiences. *Journal of Labor Economics*, 39(4):931–964.

Autor, D. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American economic review*, 103(5):1553–97.

Autor, D., Dorn, D., Hanson, G., and Majlesi, K. (2020). Importing political polarization? the electoral consequences of rising trade exposure. *American Economic Review*, 110(10):3139–3183.

Autor, D. H., Dorn, D., Hanson, G. H., and Song, J. (2014). Trade adjustment: Worker-level evidence. *The Quarterly Journal of Economics*, 129(4):1799–1860.

Avrillier, P., Hivert, L., and Kramarz, F. (2010). Driven out of employment? The impact of the abolition of national service on driving schools and aspiring drivers. *British Journal of Industrial Relations*, 48(4):784–807.

Bach, H., Campa, P., De Giorgi, G., Nosal, J., and Pietrobon, D. (2023). Born to be (sub)prime: An exploratory analysis. *CEPR Press Discussion Paper*, (17847).

Bamberger, G. E. (1987). *Occupational choice: The role of undergraduate education*. PhD thesis, The University of Chicago.

Bamieh, O. and Cintolesi, A. (2021). Intergenerational transmission in regulated professions and the role of familism. *Journal of Economic Behavior & Organization*, 192:857–879.

Barrios, J. (2018). Occupational licensing and accountant quality: Evidence from the 150-hour rule. BFI Working Papers 2018-32, Becker Friedman Institute.

Basso, G. (2009). Barriers to enter the italian lawyers' profession: An empirical surname-based approach. Master's thesis, Bocconi University.

Becker, S. O., Fetzer, T., and Novy, D. (2017). Who voted for Brexit? A comprehensive district-level analysis. *Economic Policy*, 32(92):601–650.

Beffy, M., Fougère, D., and Maurel, A. (2012). Choosing the Field of Study in Postsecondary Education: Do Expected Earnings Matter? *The Review of Economics and Statistics*, 94(1):334–347.

Bhattacharya, J. (2005). Specialty selection and lifetime returns to specialization within medicine. *Journal of Human Resources*, 40(1):115–143.

Bhuller, M. and Sigstad, H. (2022). 2sls with multiple treatments. *arXiv preprint arXiv:2205.07836*.

Bianchi, N. (2020). The indirect effects of educational expansions: Evidence from a large enrollment increase in university majors. *Journal of Labor Economics*, 38(3):767–804.

Bianchi, N. and Giorcelli, M. (2020). Scientific education and innovation: from technical diplomas to university stem degrees. *Journal of the European Economic Association*, 18(5):2608–2646.

Biasi, B. and Ma, S. (2022). The education-innovation gap. Technical report, National Bureau of Economic Research.

Bjorklund, A., Roine, J., and Waldenström, D. (2012). Intergenerational top income mobility in Sweden: Capitalist dynasties in the land of equal opportunity? *Journal of Public Economics*, 96(5):474–484.

Black, D. A., Sanders, S., and Taylor, L. (2003). The economic reward for studying economics. *Economic Inquiry*, 41(3):365–377.

Bleemer, Z. (2021). Top percent policies and the return to postsecondary selectivity. *Research & Occasional Paper Series: CSHE*, 1.

Bleemer, Z. and Mehta, A. (2022). Will studying economics make you rich? a regression discontinuity analysis of the returns to college major. *American Economic Journal: Applied Economics*, 14(2):1–22.

Bonica, A. (2014). Mapping the ideological marketplace. *American Journal of Political Science*, 58(2):367–386.

Bonica, A., McCarty, N., Poole, K. T., and Rosenthal, H. (2013). Why hasn't democracy slowed rising inequality? *Journal of Economic Perspectives*, 27(3):103–124.

Bonomi, G., Gennaioli, N., and Tabellini, G. (2021). Identity, Beliefs, and Political Conflict*. *The Quarterly Journal of Economics*, 136(4):2371–2411.

Braccioli, F., Ghinetti, P., Moriconi, S., Naguib, C., and Pellizzari, M. (2022). Education expansion, skills and labour market success.

Braga, M., Paccagnella, M., and Pellizzari, M. (2016). The impact of college teaching on students' academic and labor market outcomes. *Journal of Labor Economics*, 34(3):781–822.

Brewer, D. J., Eide, E. R., and Ehrenberg, R. G. (1999). Does it pay to attend an elite private college? cross-cohort evidence on the effects of college type on earnings. *Journal of Human Resources*, 34(1):104–123.

Brollo, F., Kaufmann, K., and La Ferrara, E. (2017). Learning about the enforcement of conditional welfare programs: Evidence from Brazil. CEPR Discussion Papers 11965, C.E.P.R. Discussion Papers.

Broscheid, A. and Teske, P. E. (2003). Public members on medical licensing boards and the choice of entry barriers. *Public Choice*, 114(3-4):445–59.

Bryson, A. and Kleiner, M. M. (2010). The regulation of occupations. *British Journal of Industrial Relations*, 48(4):670–675.

Bryson, A. and Kleiner, M. M. (2019). Re-examining advances in occupational licensing research: Issues and policy implications. *British Journal of Industrial Relations*, 57(4):721–731.

Buffington, C., Cerf, B., Jones, C., and Weinberg, B. A. (2016). Stem training and early career outcomes of female and male graduate students: Evidence from umetrics data linked to the 2010 census. *American Economic Review*, 106(5):333–38.

Buonanno, P. and Pagliero, M. (2018). Occupational licensing, labor mobility, and the unfairness of entry standards. CEPR Discussion Papers 13076, Centre for Economic Policy Research.

Buonanno, P. and Vanin, P. (2017). Social closure, surnames and crime. *Journal of Economic Behavior & Organization*, 137(C):160–175.

Bénabou, R. and Tirole, J. (2006). Incentives and prosocial behavior. *American Economic Review*, 96(5):1652–1678.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

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.

Cannings, K., Montmarquette, C., and Mahseredjian, S. (1996). Entrance quotas and admission to medical schools: A sequential probit model. *Economics of Education Review*, 15(2):163–174.

Carroll, S. L. and Gaston, R. J. (1981). Occupational restrictions and the quality of service received: Some evidence. *Southern Economic Journal*, 47(4):959–976.

Caselli, M., Fracasso, A., and Traverso, S. (2020). Globalization and electoral outcomes: Evidence from italy. *Economics & Politics*, 32(1):68–103.

Cattaneo, M. D., Jansson, M., and Ma, X. (2020). Simple local polynomial density estimators. *Journal of the American Statistical Association*, 115(531):1449–1455.

Cattaneo, M. D., Titiunik, R., and Vazquez-Bare, G. (2016). Inference in regression discontinuity designs under local randomization. *The Stata Journal*, 16(2):331–367.

Chen, M. K. and Chevalier, J. A. (2012). Are women overinvesting in education? evidence from the medical profession. *Journal of Human Capital*, 6(2):124–149.

Chernozhukov, V. and Hansen, C. (2008). The reduced form: A simple approach to inference with weak instruments. *Economics Letters*, 100(1):68–71.

Chetty, R., Friedman, J. N., Saez, E., Turner, N., and Yagan, D. (2020). Income segregation and intergenerational mobility across colleges in the United States. *Quarterly Journal of Economics*.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics*, 129(4):1553–1623.

Chise, D., Fort, M., and Monfardini, C. (2021). On the intergenerational transmission of stem education among graduate students. *The BE Journal of Economic Analysis & Policy*, 21(1):115–145.

Colantone, I. and Stanig, P. (2018). Global competition and brexit. *American Political Science Review*, 112(2):201–218.

Commission/EACEA/Eurydice, E. (2020). National student fee and support systems in european higher education – 2020/21. Technical report, Eurydice – Fact and Figures.

Corak, M. and Piraino, P. (2011). The intergenerational transmission of employers. *Journal of Labor Economics*, 29(1):37–68.

Crudu, F., Mellace, G., and Sándor, Z. (2021). Inference in instrumental variable models with heteroskedasticity and many instruments. *Econometric Theory*, 37(2):281–310.

CUN, I. N. U. C. (2000). Academic disciplines' list for italian university research and teaching.

De Giorgi, G., Drenik, A., and Seira, E. (2023). The extension of credit with nonexclusive contracts and sequential banking externalities. *American Economic Journal: Economic Policy*, 15(1):233–271.

De Giorgi, G., Harding, M., and Vasconcelos, G. F. R. (2021). Predicting mortality from credit reports. *FINANCIAL PLANNING REVIEW*, 4(4):e1135.

Delaney, J. M. and Devereux, P. J. (2019). Understanding gender differences in stem: Evidence from college applications. *Economics of Education Review*, 72:219–238.

Delaney, J. M. and Devereux, P. J. (2021). High school rank in math and english and the gender gap in stem. *Labour Economics*, 69:101969.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, 132(4):1593–1640.

Deming, D. J. and Noray, K. (2020). Earnings dynamics, changing job skills, and stem careers. *The Quarterly Journal of Economics*, 135(4):1965–2005.

Deyo, D., Hoarty, B., Norris, C., and Timmons, E. (2020). Licensing massage therapists in the name of crime: the case of Harper v Lindsay. *Journal of Entrepreneurship and Public Policy*, 10(1):1–14.

Dunn, T. and Holtz-Eakin, D. (2000). Financial capital, human capital, and the transition to self-employment: Evidence from intergenerational links. *Journal of Labor Economics*, 18(2):282–305.

d'Haultfoeuille, X. and Maurel, A. (2013). Inference on an extended Roy model, with an application to schooling decisions in France. *Journal of Econometrics*, 174(2):95–106.

Enke, B. (2020). Moral values and voting. *Journal of Political Economy*, 128(10):3679–3729.

EuroStat (2022). Tertiary education statistics.

Fariss, C. J. (2014). Respect for human rights has improved over time: Modeling the changing standard of accountability. *American Political Science Review*, 108(2):297–318.

Farronato, C., Fradkin, A., Larsen, B., and Brynjolfsson, E. (2020). Consumer protection in an online world: An analysis of occupational licensing. Working Paper 26601, National Bureau of Economic Research.

Foote, C. L., Loewenstein, L., and Willen, P. S. (2021). Cross-sectional patterns of mortgage debt during the housing boom: evidence and implications. *The Review of Economic Studies*, 88(1):229–259.

Fricke, H., Grogger, J., and Steinmayr, A. (2018). Exposure to academic fields and college major choice. *Economics of Education Review*, 64:199–213.

Friedman, M. and Kuznets, S. (1945). *Income from Independent Professional Practice*. National Bureau of Economic Research, Inc.

Funke, M., Schularick, M., and Trebesch, C. (2016). Going to extremes: Politics after financial crises, 1870–2014. *European Economic Review*, 88:227–260.

Granato, S. (2018). Gender inequalities and scarring effects in school to work transitions.

Grossman, G. M. and Helpman, E. (2020). Identity Politics and Trade Policy. *The Review of Economic Studies*, 88(3):1101–1126.

Güell, M., Mora, J. V. R., and Telmer, C. I. (2015). The Informational Content of Surnames, the Evolution of Intergenerational Mobility, and Assortative Mating. *Review of Economic Studies*, 82(2):693–735.

Güell, M., Pellizzari, M., Pica, G., and Mora, J. V. R. (2018). Correlating social mobility and economic outcomes. *Economic Journal*, 128(612):F353–F403.

Guiso, L., Herrera, H., Morelli, M., and Sonno, T. (2017). Demand and Supply of Populism. EIEF Working Papers Series 1703, Einaudi Institute for Economics and Finance (EIEF).

Guriev, S. and Papaioannou, E. (2022). The political economy of populism. *Journal of Economic Literature*, 60(3):753–832.

Haas-Wilson, D. (1986). The effect of commercial practice restrictions: The case of optometry. *The Journal of Law and Economics*, 29(1):165–186.

Hall, A. B. and Yoder, J. (2022). Does homeownership influence political behavior? Evidence from administrative data. *The Journal of Politics*, 84(1):351–366.

Hanson, M. (2022). College enrollment & student demographic statistics.

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

Heckman, J. J. and Mosso, S. (2014). The economics of human development and social mobility. *Annual Review of Economics*, 6(1):689–733.

Heckman, J. J. and Pinto, R. (2018). Unordered monotonicity. *Econometrica*, 86(1):1–35.

Heckman, J. J. and Urzua, S. (2010). Comparing IV with structural models: What simple IV can and cannot identify. *Journal of Econometrics*, 156(1):27–37.

Heckman, J. J., Urzua, S., and Vytlacil, E. (2006). Understanding instrumental variables in models with essential heterogeneity. *The Review of Economics and Statistics*, 88(3):389–432.

Herkenhoff, K., Phillips, G., and Cohen-Cole, E. (2016). How credit constraints impact job finding rates, sorting & aggregate output. Working Paper 22274, National Bureau of Economic Research.

Herkenhoff, K., Phillips, G. M., and Cohen-Cole, E. (2021). The impact of consumer credit access on self-employment and entrepreneurship. *Journal of Financial Economics*, 141(1):345–371.

Hoffman, S. D. and Duncan, G. J. (1988). Multinomial and conditional logit discrete-choice models in demography. *Demography*, 25(3):415–427.

Hotz, V. J. and Miller, R. A. (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies*, 60(3):497–529.

ISTAT (2017). *Definizione dei gruppi sociali e loro descrizione.* Note to ch. 2, Rapporto Annuale ISTAT.

ISTAT (2021). Iscritti all'università. `http://dati.istat.it/Index.aspx?DataSetCode=DCIS_ISCRITTI`.

Kahn, S. and Ginther, D. (2017). Women and stem. Technical report, National Bureau of Economic Research.

Kara, G. and Yook, Y. (2022). Policy uncertainty and bank mortgage credit. *Journal of Money, Credit and Banking*, forthcoming.

Ketel, N., Leuven, E., Oosterbeek, H., and van der Klaauw, B. (2016). The returns to medical school: Evidence from admission lotteries. *American Economic Journal: Applied Economics*, 8(2):225–54.

Kirkeboen, L. J., Leuven, E., and Mogstad, M. (2016). Field of study, earnings, and self-selection. *The Quarterly Journal of Economics*, 131(3):1057–1111.

Kleiner, M. M. (2000). Occupational licensing. *Journal of Economic Perspectives*, 14(4):189–202.

Kleiner, M. M. (2017). The influence of occupational licensing and regulation. *IZA World of Labor*, pages 392–392.

Kleiner, M. M. and Krueger, A. B. (2013). Analyzing the extent and influence of occupational licensing on the labor market. *Journal of Labor Economics*, 31(S1):173–202.

Kleiner, M. M. and Kudrle, R. T. (2000). Does regulation affect economic outcomes? The case of dentistry. *The Journal of Law and Economics*, 43(2):547–582.

Kleiner, M. M., Marier, A., Park, K. W., and Wing, C. (2016). Relaxing occupational licensing requirements: Analyzing wages and prices for a medical service. *The Journal of Law and Economics*, 59(2):261–291.

Kleiner, M. M. and Soltas, E. J. (2022). A welfare analysis of occupational licensing in US states. *Review of Economics Studies*, Forthcoming.

Koumenta, M. and Pagliero, M. (2018). Occupational regulation in the european union: Coverage and wage effects. *British Journal of Industrial Relations*, 57(4):818–849.

Kugler, A. D. and Sauer, R. M. (2005). Doctors without borders? Relicensing requirements and negative selection in the market for physicians. *Journal of Labor Economics*, 23(3):437–465.

Laband, D. and Lentz, B. F. (1992). Self-recruitment in the legal profession. *Journal of Labor Economics*, 10(2):182–201.

Larsen, B., Ju, Z., Kapor, A., and Yu, C. (2020). The effect of occupational licensing stringency on the teacher quality distribution. NBER Working Papers 28158, National Bureau of Economic Research, Inc.

Lee, D. and van der Klaauw, W. (2010). Introduction to the FRBNY Consumer Credit Panel. *FRBNY Staff Report*, (479).

Leip, D. (2017). David leip's atlas of u.s. presidential elections, datasets. *Harvard Dataverse, https://doi. org/10.7910/DVN/XX3YJ4*.

Leland, H. E. (1979). Quacks, lemons, and licensing: A theory of minimum quality standards. *Journal of political economy*, 87(6):1328–1346.

Lentz, B. F. and Laband, D. (1989). Why so many children of doctors become doctors: Nepotism vs. human capital transfers. *Journal of Human Resources*, 24(3):396–413.

Maddala, G. S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*. Econometric Society Monographs. Cambridge University Press.

Malamud, O. (2010). Breadth versus depth: the timing of specialization in higher education. *Labour*, 24(4):359–390.

Malamud, O. (2011). Discovering one's talent: learning from academic specialization. *ILR Review*, 64(2):375–405.

Mann, A. and DiPrete, T. A. (2013). Trends in gender segregation in the choice of science and engineering majors. *Social Science Research*, 42(6):1519–1541.

Maple, S. A. and Stage, F. K. (1991). Influences on the choice of math/science major by gender and ethnicity. *American Educational Research Journal*, 28(1):37–60.

Matsusaka, J. G. (1995). Fiscal effects of the voter initiative: Evidence from the last 30 years. *Journal of political Economy*, 103(3):587–623.

Maurizi, A. (1974). Occupational licensing and the public interest. *The Journal of Political Economy*, 82(2):399–413.

McCarty, N., Poole, K. T., and Rosenthal, H. (2016). *Polarized America: The dance of ideology and unequal riches.*

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2):698–714.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*, pages 105–142.

McFadden, D. L. (1984). Econometric analysis of qualitative response models. In *Handbook of Econometrics*, volume 2, pages 1395–1457. Elsevier.

Mian, A. and Sufi, A. (2011). House prices, home equity–based borrowing, and the US household leverage crisis. *American Economic Review*, 101(5):2132–2156.

Mian, A., Sufi, A., and Trebbi, F. (2010). The Political Economy of the US Mortgage Default Crisis. *American Economic Review*, 100(5):1967–98.

Mikusheva, A. and Sun, L. (2022). Inference with many weak instruments. *The Review of Economic Studies*, 89(5):2663–2686.

Mocetti, S. (2016). Dynasties in professions and the role of rents and regulation: Evidence from Italian pharmacies. *Journal of Public Economics*, 133:1–10.

Mocetti, S. and Roma, G. (2021). Le professioni ordinistiche in Italia: misure ed effetti della regolamentazione. *Mercato Concorrenza Regole*, 3:611–639.

Mocetti, S., Roma, G., and Rubolino, E. (2022). Knocking on parents' doors: Regulation and intergenerational mobility. *Journal of Human Resources*, 52:525–554.

Montmarquette, C., Cannings, K., and Mahseredjian, S. (2002). How do young people choose college majors? *Economics of Education Review*, 21(6):543–556.

Mountjoy, J. (2022). Community colleges and upward mobility. *forthcoming American Economic Review.*

Mutz, D. C. (2018). Status threat, not economic hardship, explains the 2016 presidential vote. *Proceedings of the National Academy of Sciences*, 115(19):E4330–E4339.

National Credit Union Administration (2016). 2004-2016 annual report. Technical report.

Neal, D. A. (2018). *Information, incentives, and education policy*. Harvard University Press.

Ng, K. and Riehl, E. (2020). The returns to stem programs for less-prepared students.

Office of Small Credit Union Initiatives (2016). Serving the credit invisibles. Technical report.

Oreopoulos, P. and Petronijevic, U. (2013). Making college worth it: A review of the returns to higher education. *Future of Children*, 23(1):41–65.

Orsini, J. and Pellizzari, M. (2012). *Dinastie d'Italia: gli ordini tutelano davvero i consumatori?* Università Bocconi Editore.

Pagliero, M. (2010). Licensing exam difficulty and entry salaries in the us market for lawyers. *British Journal of Industrial Relations*, 48(4):726–739.

Pagliero, M. (2011). What is the objective of professional licensing? evidence from the us market for lawyers. *International Journal of Industrial organization*, 29(4):473–483.

Paterson, I., Fink, M., and Ogus, A. (2003). *Economic impact of regulation in the field of liberal professions in different Member States*. Institute for Advanced Studies (IHS).

Patnaik, A., Wiswall, M., and Zafar, B. (2020). College majors. *NBER Working Paper*, (w27645).

Pellizzari, M., Basso, G., Catania, A., Labartino, G., Malacrino, D., and Monti, P. (2011). Family ties in licensed professions in Italy. Technical report, Fondazione Rodolfo Debenedetti.

Pellizzari, M. and Pica, G. (2010). Liberalizing professional services: Evidence from Italian lawyers. Technical Report 372, Bocconi University.

Persson, T. and Tabellini, G. (2002). Political economics and public finance. *Handbook of public economics*, 3:1549–1659.

Phillips, P. C. and Gao, W. Y. (2017). Structural inference from reduced forms with many instruments. *Journal of Econometrics*, 199(2):96–116.

Poole, K. T. and Rosenthal, H. (1985). A spatial model for legislative roll call analysis. *American journal of political science*, pages 357–384.

Poole, K. T. and Rosenthal, H. (1991). Patterns of congressional voting. *American journal of political science*, pages 228–278.

Porter, C. and Serra, D. (2020). Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics*, 12(3):226–54.

Porter, J. and Yu, P. (2015). Regression discontinuity designs with unknown discontinuity points: Testing and estimation. *Journal of Econometrics*, 189(1):132–147.

Raitano, M. and Vona, F. (2021). Nepotism vs specific skills : the effect of professional liberalization on returns to parental back grounds of Italian lawyers. *Journal of Economic Behavior & Organization*, 184:489–505.

Rask, K. (2010). Attrition in stem fields at a liberal arts college: The importance of grades and pre-collegiate preferences. *Economics of Education Review*, 29(6):892–900.

Rodrik, D. and Mukand, S. (2018). The political economy of ideas: On ideas versus interests in policymaking. Technical report.

Schmeiser, M., Stoddard, C., and Urban, C. (2016). Student loan information provision and academic choices. *American Economic Review*, 106(5):324–28.

Shapiro, C. (1986). Investment, moral hazard, and occupational licensing. *The Review of Economic Studies*, 53(5):843–862.

Stigler, G. J. (1971). The Theory of Economic Regulation. *Bell Journal of Economics*, 2(1):3–21.

Thornton, R. J. and Weintraub, A. R. (1979). Licensing in the barbering profession. *ILR Review*, 32(2):242–249.

Timmons, E. J. and Thornton, R. J. (2008). The effects of licensing on the wages of radiologic technologists. *Journal of Labor Research*, 29(4):333–346.

Timmons, E. J. and Thornton, R. J. (2010). The licensing of barbers in the USA. *British Journal of Industrial Relations*, 48(4):740–757.

Uddin, S., Imam, T., and Mozumdar, M. (2021). Research interdisciplinarity: Stem versus non-stem. *Scientometrics*, 126(1):603–618.

UK Office of Fair Trade (2001). *Competition in Professions.* OFT.

US Census Bureau (2019). Number of people with master's and doctoral degrees doubles since 2000. https://www.census.gov/library/stories/2019/02/number-of-people-with-masters-and-phd-degrees-double-since-2000.html.

Vytlacil, E. (2002). Independence, monotonicity, and latent index models: An equivalence result. *Econometrica*, 70(1):331–341.

Wanchek, T. (2010). Dental hygiene regulation and access to oral healthcare: Assessing the variation across the US states. *British Journal of Industrial Relations*, 48(4):706–725.

Webber, D. A. (2016). Are college costs worth it? How ability, major, and debt affect the returns to schooling. *Economics of Education Review*, 53:296–310.

Wing, P., Langelier, M. H., Continelli, T. A., and Battrell, A. (2005). A dental hygiene professional practice index (DHPPI) and access to oral health status and service use in the united states. *Journal of Dental Hygiene*, 79(2).

Winters, J. V. (2014). Stem graduates, human capital externalities, and wages in the us. *Regional Science and Urban Economics*, 48:190–198.

Xie, Y., Fang, M., and Shauman, K. (2015). STEM education. *Annual review of sociology*, 41:331–357.

Appendices

APPENDIX A

Appendix to Chapter 1: A model of professional human capital accumulation

In this appendix, we sketch a simple two-period model of human capital accumulation. For expositional clarity, we refer to the two periods as law school and apprentices but they could really be anything else, such as the family and law school.

We assume that individuals (with a certain predetermined level of general human capital A) first accumulate occupation-specific human capital at law school and then during the apprenticeship. At each step, human capital accumulation requires effort and depends positively on the stock accumulated up to that step.

We assume that the production function of occupation-specific human capital at law school (that we denote as GPA) takes the following form:

$$GPA = \Phi_1 A^{\beta_1} e_1^{\beta_2}$$

where e_1 denotes the amount of endogenous effort exerted in period 1. A and Φ_1 are positive parameters. The former reflects the fact that higher general human capital favours the accumulation of occupation-specific human capital; the latter parametrizes the effectiveness of the learning process at law school, a factor which is fully captured in our empirical specification by university fixed effects.

During the apprenticeship period (period 2) there is further accumulation of occupation-specific human capital S_2. We assume that:

$$S_2 = \Phi_2 A^{\gamma_1} (GPA)^{\gamma_2} e_2^{\gamma_3} \tag{A.1}$$

where e_2 denotes the amount of endogenous effort exerted in period 2 and Φ_2 is a positive parameter that reflects the quality of the law firm where the apprenticeship takes place. The amount of general and occupation-specific human capital previously accumulated (i.e. A and GPA) also contributes to the process. We assume that there are decreasing returns to scale in overall effort, i.e. $\gamma_2\beta_2 + \gamma_3 < 1$.

Finally, lifetime earnings are an increasing and concave function of the total amount of occupation-specific human capital S_2:

$$w = f(S_2)$$

We assume that individuals choose effort in each period to maximize lifetime earnings and that exerting effort is costly. For simplicity, we assume that $c(e_t) = \delta_t e_t$ for $t = 1, 2$ (any increasing and convex cost function would generate similar results). As a result, the individual solves the following problem:

$$\begin{aligned} & \max_{e_1,e_2} f(S_2) - \delta_1 e_1 - \delta_2 e_2 \\ & s.t. \\ GPA &= \Phi_1 A^{\beta_1} e_1^{\beta_2} \\ S_2 &= \Phi_2 A^{\gamma_1} (GPA)^{\gamma_2} e_2^{\gamma_3} \end{aligned}$$

The first-order conditions are:

$$\beta_2\gamma_2 f'(S_2)\Phi_2 A^{(\gamma_1+\beta_1\gamma_2)}\Phi_1^{\gamma_2}\left[e_1^{\beta_2\gamma_2-1}e_2^{\gamma_3}\right] = \delta_1 \tag{A.2}$$

$$\gamma_3 f'(S_2)\Phi_2 A^{(\gamma_1+\beta_1\gamma_2)}\Phi_1^{\gamma_2}\left[e_1^{\beta_2\gamma_2}e_2^{\gamma_3-1}\right] = \delta_2 \tag{A.3}$$

Notice that the marginal productivity of effort in each period depends positively on the effort exerted in the other period. Dividing equation (A.2) by equation (A.3) we get:

$$\frac{e_2}{e_1} = \frac{\delta_1}{\delta_2}\frac{\gamma_3}{\beta_2\gamma_2} \tag{A.4}$$

Notice the following:

1. Neither the Φs, nor A, nor the shape of $f(\cdot)$ affect relative effort between periods. This implies that it is never the case that the individual wants to make little effort in period 1, during law school, and a lot of effort in period 2, during the apprenticeship period, because the human capital accumulation process is more effective during the apprenticeship than at law school, i.e. because $\Phi_2 > \Phi_1$. Complementarity between GPA and e_2 in the production function of human capital in equation (A.1) implies that the agent does want

to exert effort in period 1 even if Φ_1 is low relative to Φ_2 because this raises the marginal productivity of effort in period 2.

2. What matters in shaping relative effort is the relative marginal cost (δs) and the curvature of the human capital production function. The more concave the human capital production function in a specific period, the lower effort in that period.

Substituting back equation (A.4) into equation (A.2) we get:

$$\beta_2\gamma_2 f'(S_2)\Phi_2 A^{(\gamma_1+\beta_1\gamma_2)}\Phi_1^{\gamma_2}\left[e_1^{\beta_2\gamma_2-1}e_2^{\gamma_3}\right] = \delta_1 \tag{A.5}$$

$$\beta_2\gamma_2 f'(S_2)\Phi_2 A^{(\gamma_1+\beta_1\gamma_2)}\Phi_1^{\gamma_2}\left[e_1^{\beta_2\gamma_2-1}\left(e_1\frac{\delta_1}{\delta_2}\frac{\gamma_3}{\beta_2\gamma_2}\right)^{\gamma_3}\right] = \delta_1 \tag{A.6}$$

$$\beta_2\gamma_2 f'(S_2)\Phi_2 A^{(\gamma_1+\beta_1\gamma_2)}\Phi_1^{\gamma_2}\left[e_1^{\beta_2\gamma_2-1+\gamma_3}\left(\frac{\delta_1}{\delta_2}\frac{\gamma_3}{\beta_2\gamma_2}\right)^{\gamma_3}\right] = \delta_1 \tag{A.7}$$

The above equation implicitly defines the optimal amount of effort in period 1.

Assuming that $f(\cdot)$ is linear we get a closed-form solution (qualitatively, results hold with any $f(\cdot)$, provided that $f(\cdot)$ is not too concave):

$$e_1^* = \left[\frac{\beta_2\gamma_2\Phi_2 A^{(\gamma_1+\beta_1\gamma_2)}\Phi_1^{\gamma_2}\left(\frac{\delta_1}{\delta_2}\frac{\gamma_3}{\beta_2\gamma_2}\right)^{\gamma_3}}{\delta_1}\right]^{\frac{1}{1-\beta_2\gamma_2-\gamma_3}} \tag{A.8}$$

$$e_2^* = e_1^*\frac{\delta_1}{\delta_2}\frac{\gamma_3}{\beta_2\gamma_2} \tag{A.9}$$

The model shows that – holding fixed general human capital and the quality of the law school – a higher expected quality of the law firm (Φ_2) raises e_1^* and therefore also the amount of occupation-specific human capital accumulated at law school (GPA). Note that the quality of the law firm in our context captures also the ability of senior lawyers in transferring soft skills to apprentices. Thus, candidates who are connected to lawyers who can transfer them soft skills are still incentivized to exert more effort during law school to get a higher GPA.

Of course, this result relies on the assumption that the amount of occupation-specific human capital accumulated at law school and effort during the apprenticeship are complements in the production of further occupation-specific human capital. Should instead the two inputs be substitutes, this result would be overturned. However, we view complementarity as the most natural and plausible assumption.

This result would be reversed, for instance, if the production function of human capital featured a high degree of substitutability between the human capital accumulated in law school and the effort exerted during the apprenticeship.[1] We believe that such a situation is highly unlikely, but it could be generated, for example, by connected individuals having access to better training firms during their apprenticeships. In this scenario, connected agents might choose to exert little effort in law school and then, thanks to their connections, recoup the lost human capital through apprenticeships in high-quality law firms. In this setting, GPA in law school might even be negatively correlated with professional ability at the end of the apprenticeship.

As a confirmation that this scenario does not square with empirical evidence, in Section 1.5.4 we showed that there is no evidence of connected individuals receiving additional training from high-quality lawyers during their apprenticeship.

[1] Notice that a small degree of substitutability would not suffice to undo the positive relationship between law school performance and final professional ability.

APPENDIX B

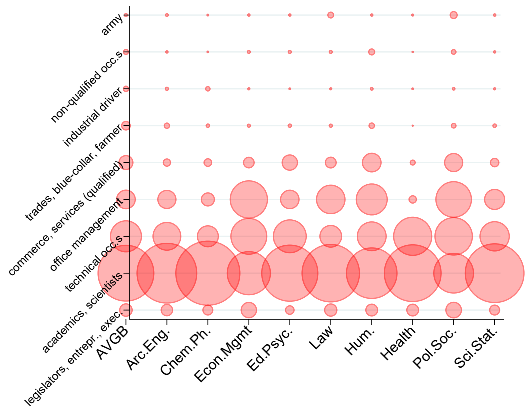
Appendix to Chapter 2: Additional Descriptives and Results, Methodological Notes and Definitions

B.1 Additional Descriptive Results

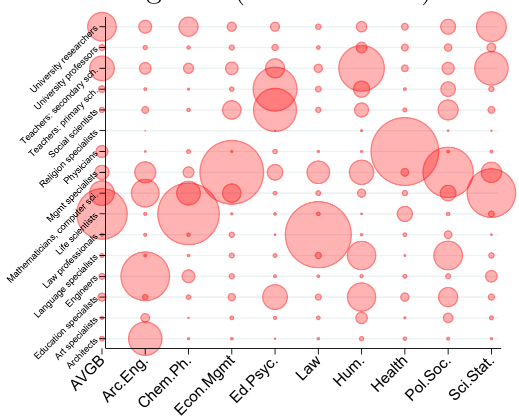
Table B.1: Differences in X for the sample of employed and unemployed.

	All	Employed	Unemployed
High School: grade (st.)	0.00	0.03	-0.11
	(1.000)	(0.998)	(0.998)
High School: humanities	0.15	0.14	0.18
	(0.359)	(0.352)	(0.384)
High School: science	0.39	0.40	0.36
	(0.487)	(0.489)	(0.479)
Gender (1=female)	0.62	0.60	0.68
	(0.485)	(0.489)	(0.466)
Parents: graduate	0.26	0.26	0.26
	(0.438)	(0.437)	(0.439)
Parents: high-ranked occ.	0.21	0.22	0.21
	(0.410)	(0.412)	(0.406)
Employment	0.77	1.00	0.00
	(0.418)	(0)	(0)
Observations	655 847	508 242	147 605

Figure B.1: Occupation sectors by master degree's (ISTAT codes)



(a) Share of master graduates in one-digit occupations.



(b) Two-digit occupations for master graduates employed in sector "academics and scientists".

Panel B.1a presents one-digit occupation sectors for all master graduates as defined by ISTAT's 2011 classification of occupations (in turn based on ILO's 2008 *International Standard Classification of Occupations*). This information is available for 209 906 individuals. Panel B.1b focuses on two-digit occupation sectors for master graduates employed in sector "academics and scientists" (intellectual and highly specialized occupations), for a total of 126 166 observations. In both instances, occupation codes are only available for individuals who complete a master degree and are not available for students who start working after their bachelor. Both panels show that labor markets are segregated along specialized skill sets.

Table B.2: Comparison of STEM classification methods in Economics papers

Paper	**Classification**	**Groups of Degrees Classified as STEM**									
		Science	**Arch. Eng.**	**Chem. Pharm.**	**Agr.Vet. Geo.Bio.**	**Econ. Stat.**	**Health**	**Educ. Psy.**	**Law**	**Lit. Lang.**	**Pol. Soc.**
Adams and Kirchmaier (2016)	O*NET, authors	All	Most	Most	Some	Most	None	None	None	None	None
Ahn et al. (2019)	author	All	Most	Most	Some	All	None	None	None	None	None
Altonji et al. (2016)	author	All	All	All	All	All	All	None	None	None	None
Altonji and Zhong (2021)	author, NSCG, NSRCG	All	Most	None	All	None	None	None	None	None	None
Arcidiacono et al. (2016b)	author	All	All	All	Most	Some	None	Some	None	None	None
Arcidiacono et al. (2016a)	author	All	All	Most	All	None	Some	None	None	None	None
Bianchi and Giorcelli (2020)	author	All	Most	Most	Most	Some	None	None	None	None	None
Biasi and Ma (2022)	ARC 2016	All	All	All	All	None	All	None	None	None	None
Buffington et al. (2016)	author	All	Most	All	Most	None	All	None	None	None	None
Canaan and Mouganie (2018)	author	All	All	All	All	Most	Most	None	None	None	None
Chise et al. (2021)	ISCED, MIUR, author	All	All	Most	Some	Some	None	None	None	None	None
Delaney and Devereux (2019)	ISCED, authors	All	All	All	Some	Some	All	None	None	None	None
Delaney and Devereux (2021)	ISCED, authors	All	All	All	Some	Some	All	None	None	None	None
Deming (2017)	Autor and Dorn (2013)	All	All	Some	Some	Some	Some	None	None	None	None
Granato (2018)	MIUR	All	All	All	Some	Some	None	None	None	None	None
Kahn and Ginther (2017)	author	All	Most	All	Most	Most	None	Some	None	None	None
Maple and Stage (1991)	author	All	Most	None	None	None	None	None	None	None	None
Ng and Riehl (2020)	author	All	Most	Most	Most	Some	None	None	None	None	None
Porter and Serra (2020)	author	All	All	All	None	All	None	None	None	None	None
Rask (2010)	Anon. data provider	All	NA	All	Most	Some	None	Some	None	None	None
Schmeiser et al. (2016)	author	All	All	All	All	Some	Most	None	None	None	None
Uddin et al. (2021)	ARC 2016	All	All	All	All	None	All	None	None	None	None
Webber (2016)	author, NLSY, NSCG, ACS	All	Most	Most	Some	None	None	None	None	None	None
Winters (2014)	U.S. ICE	All	All	All	All	Some	Some	Some	None	None	None

The table presents a non-exhaustive review of the methods used to classify STEM fields in the literature. All possible degrees are grouped into ten categories: Science, Architecture and Engineering, Chemistry and Pharmacy, Agriculture Veterinary Geology and Biology, Economics and Statistics, Education and Psychology, Law, Literature and Languages, Political sciences and social sciences. The full list of degrees belonging to each group according to the Italian classification can be found in appendix B.5.2. The grouping and the comparison across papers is inherently lax as not all degrees are available across countries. The label "Most" indicates that almost all the degrees in the group according to my grouping are defined as STEM in the authors' paper. "Some" indicates that only a few of them are defined as STEM. "All" and "None" should be self-explanatory.

B.2 Marginal Effects of Main Regressions

Table B.3: $t = 1$: Marginal Effects at Means of exclusion restrictions on choice of bachelor

Z_j: *Entry Exams*	AVGB (1)	Ar.Eng. (2)	Ch.Pharm. (3)	Ec.Mgmt. (4)	Ed.Psy. (5)	Law (6)	Lit.Lan. (7)	Health (8)	Pol.Soc. (9)	Sci.Stat. (10)
Outcomes										
Pr(AVGB)	-0.031***	0.004	0.048***	0.016***	-0.014***	0.049***	-0.170***	0.196***	-0.065***	0.046***
	(0.003)	(0.004)	(0.002)	(0.002)	(0.002)	(0.003)	(0.005)	(0.007)	(0.005)	(0.004)
Pr(ArEng)	-0.061***	0.053***	0.138***	0.025***	0.049***	0.046***	0.173***	-0.127***	-0.280***	-0.022***
	(0.004)	(0.005)	(0.003)	(0.002)	(0.003)	(0.004)	(0.006)	(0.009)	(0.007)	(0.006)
Pr(ChPh)	-0.005**	0.044***	-0.009***	-0.011***	-0.026***	0.000	-0.043***	0.137***	0.041***	0.016***
	(0.002)	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)	(0.005)	(0.004)	(0.003)
Pr(EcMg)	-0.171***	-0.141***	0.064***	0.013***	-0.031***	0.026***	0.305***	0.080***	0.046***	-0.047***
	(0.004)	(0.005)	(0.003)	(0.002)	(0.003)	(0.004)	(0.006)	(0.009)	(0.006)	(0.005)
Pr(EdPsy)	0.091***	-0.097***	-0.042***	0.021***	0.035***	-0.025***	-0.145***	0.042***	-0.073***	0.125***
	(0.003)	(0.004)	(0.003)	(0.002)	(0.003)	(0.004)	(0.006)	(0.008)	(0.006)	(0.005)
Pr(Law)	0.049***	-0.036***	0.018***	0.005**	-0.059***	0.018***	-0.050***	-0.056***	0.003	0.044***
	(0.003)	(0.004)	(0.003)	(0.002)	(0.002)	(0.003)	(0.005)	(0.008)	(0.006)	(0.005)
Pr(LitLan)	0.003	-0.037***	0.076***	0.055***	-0.085***	-0.156***	0.264***	-0.543***	0.034***	-0.016***
	(0.004)	(0.005)	(0.003)	(0.003)	(0.003)	(0.004)	(0.006)	(0.010)	(0.007)	(0.006)
Pr(Health)	0.078***	0.168***	-0.286***	-0.121***	0.184***	0.119***	-0.302***	0.462***	0.171***	-0.182***
	(0.004)	(0.005)	(0.003)	(0.002)	(0.003)	(0.004)	(0.007)	(0.010)	(0.007)	(0.006)
Pr(PolSoc)	0.037***	0.014***	0.020***	-0.009***	-0.063***	-0.067***	-0.035***	-0.210***	0.104***	0.038***
	(0.003)	(0.005)	(0.003)	(0.002)	(0.003)	(0.004)	(0.006)	(0.009)	(0.007)	(0.005)
Pr(SciSt)	0.010***	0.026***	-0.027***	0.007***	0.008***	-0.009***	0.003	0.019***	0.017***	-0.002
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.005)	(0.003)	(0.003)
Observations	655,847	655,847	655,847	655,847	655,847	655,847	655,847	655,847	655,847	655,847

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

B.3 Choice of Master Degree – Regression Tables

Table B.4: Choice of master conditional on bachelor in Agriculture, Veterinary, Geology, Biology

VARIABLES	(1) AVGB	(2) Arc.Eng.	(3) Ed.Psy.	(4) Health	(5) Sci.Stat.
Z_m					
Credit req. (st.): AVGB	1.141***	-1.191**	1.727	-6.138***	6.461***
	(0.114)	(0.530)	(1.472)	(0.777)	(0.251)
Credit req. (st.): ArEn	3.897***	2.266*	13.779***	17.824***	1.226
	(0.323)	(1.242)	(4.057)	(1.164)	(1.031)
Credit req. (st.): Med	-2.667***	7.017***	-8.738***	2.887***	-1.746***
	(0.132)	(0.932)	(1.467)	(0.472)	(0.450)
Credit req. (st.): Sci	-1.668***	0.931	-15.627***	-6.154***	-9.848***
	(0.140)	(0.863)	(3.481)	(0.596)	(1.024)
log(distance)	0.006	-0.166***	0.043	-0.012	-0.016
	(0.007)	(0.029)	(0.041)	(0.021)	(0.018)
X					
HS: grade (st.)	0.538***	0.237***	0.124	0.314***	0.454***
	(0.016)	(0.076)	(0.090)	(0.048)	(0.041)
HS: humanities	0.872***	0.402	0.638***	0.547***	1.121***
	(0.055)	(0.277)	(0.225)	(0.141)	(0.135)
HS: science	0.867***	0.140	-0.390**	0.488***	0.860***
	(0.032)	(0.163)	(0.191)	(0.102)	(0.088)
Gender (1=female)	-0.451***	-0.097	0.667***	0.077	0.252**
	(0.041)	(0.205)	(0.252)	(0.134)	(0.107)
Parents: graduate	0.366***	0.387**	0.361*	0.406***	0.468***
	(0.040)	(0.185)	(0.202)	(0.114)	(0.095)
Parents: high-rank occ.	-0.033	0.198	0.209	0.089	-0.126
	(0.042)	(0.191)	(0.208)	(0.124)	(0.106)
Additional Controls			Yes		
Θ			Yes		
Constant	10.364***	-34.833***	17.671**	-41.140***	13.965***
	(0.909)	(5.033)	(7.548)	(3.938)	(2.570)
Observations	32,494	32,494	32,494	32,494	32,494
Pseudo R2	0.211	0.211	0.211	0.211	0.211

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Excluded category: no master.

Table B.5: Choice of master conditional on bachelor in Architecture and Engineering

VARIABLES	(1) Arc.Eng.	(2) Chem.Pharm.	(3) Lit.Lang.	(4) Sci.Stat.
Z_m				
Credit req. (st.): ArEn	-0.053**	-2.726***	-6.861***	-0.268*
	(0.024)	(0.144)	(0.687)	(0.143)
Credit req. (st.): ChPh	-6.053***	-1.743*	23.136***	-5.426***
	(0.127)	(0.977)	(2.300)	(1.132)
Credit req. (st.): Hum	-0.734***	15.535***	14.477***	8.470***
	(0.127)	(2.934)	(2.252)	(2.453)
Credit req. (st.): Sci	0.649***	0.239	5.577***	-1.262*
	(0.053)	(0.729)	(0.521)	(0.754)
log(distance)	-0.015***	0.005	0.004	0.005
	(0.005)	(0.022)	(0.028)	(0.037)
X				
HS: grade (st.)	0.737***	0.742***	0.634***	0.808***
	(0.010)	(0.045)	(0.072)	(0.079)
HS: humanities	0.900***	1.340***	1.124***	0.849***
	(0.044)	(0.170)	(0.192)	(0.302)
HS: science	1.023***	1.425***	0.564***	1.273***
	(0.021)	(0.098)	(0.148)	(0.164)
Gender (1=female)	-0.253***	0.208*	0.549***	0.604***
	(0.030)	(0.113)	(0.194)	(0.179)
Parents: graduate	0.368***	0.385***	0.517***	0.544***
	(0.025)	(0.089)	(0.151)	(0.150)
Parents: high-rank occ.	0.074***	-0.033	-0.029	0.015
	(0.026)	(0.097)	(0.160)	(0.162)
Additional Controls		Yes		
Θ		Yes		
Constant	4.078***	-2.262	-18.458***	-0.826
	(0.158)	(1.656)	(1.929)	(1.670)
Observations	79,817	79,817	79,817	79,817
Pseudo R2	0.241	0.241	0.241	0.241

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Excluded category: no master.

Table B.6: Probability of choosing a master degree given a bachelor in Chemistry and Pharmacy

VARIABLES	(1) AVGB	(2) Chem.Pharm.	(3) Health
Z_m			
Credit req. (st.): ChPh	0.423***	-2.994***	7.038
	(0.125)	(0.186)	(293.390)
log(distance)	-0.072	-0.013	-0.037
	(0.044)	(0.021)	(0.037)
X			
HS: grade (st.)	0.688***	0.770***	0.779***
	(0.112)	(0.047)	(0.088)
HS: humanities	0.346	0.877***	0.864***
	(0.315)	(0.176)	(0.239)
HS: science	-0.049	0.866***	0.622***
	(0.217)	(0.093)	(0.167)
Gender (1=female)	0.300	-0.203	-0.781***
	(0.283)	(0.126)	(0.244)
Parents: graduate	0.464*	0.403***	-0.013
	(0.248)	(0.115)	(0.230)
Parents: high-rank occ.	0.197	0.080	-0.002
	(0.272)	(0.130)	(0.261)
Additional Controls		Yes	
Θ		Yes	
Constant	-3.444***	-5.804***	-23.672
	(0.911)	(0.580)	(1,430.729)
Observations	7,398	7,398	7,398
Pseudo R2	0.571	0.571	0.571

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Excluded category: no master.

Table B.7: Probability of choosing a master degree given a bachelor in Economics and Management

VARIABLES	(1) Econ.Mgmt.	(2) Educ.Psy.	(3) Law	(4) Pol.Soc.	(5) Sci.Stat.
Z_m					
Credit req. (st.): PlSc	-18.471***	-15.756***	-13.175***	-18.097***	-13.711***
	(0.157)	(1.497)	(1.264)	(0.554)	(0.815)
log(distance)	0.012**	-0.125***	0.040**	-0.054***	-0.025
	(0.005)	(0.046)	(0.019)	(0.016)	(0.028)
X					
HS: grade (st.)	0.509***	0.027	-0.009	0.237***	0.515***
	(0.010)	(0.101)	(0.075)	(0.034)	(0.053)
HS: humanities	0.908***	1.391***	1.506***	1.160***	0.875***
	(0.041)	(0.282)	(0.211)	(0.105)	(0.206)
HS: science	0.763***	0.771***	-0.160	0.538***	1.128***
	(0.021)	(0.213)	(0.182)	(0.072)	(0.106)
Gender (1=female)	-0.393***	1.282***	-1.303***	-0.065	-0.432***
	(0.026)	(0.258)	(0.221)	(0.088)	(0.133)
Parents: graduate	0.377***	0.698***	-0.091	0.608***	0.442***
	(0.026)	(0.225)	(0.216)	(0.080)	(0.122)
Parents: high-rank occ.	0.076***	-0.392	-0.868***	-0.105	-0.194
	(0.025)	(0.248)	(0.261)	(0.085)	(0.130)
Additional Controls			Yes		
Θ			Yes		
Constant	-40.752***	-44.205***	-31.250***	-41.948***	-34.294***
	(0.365)	(3.554)	(2.950)	(1.284)	(1.900)
Observations	75,993	75,993	75,993	75,993	75,993
Pseudo R2	0.220	0.220	0.220	0.220	0.220

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Excluded category: no master.

Table B.8: Probability of choosing a master degree given a bachelor in Physical Education, Teaching and Psychology

VARIABLES	(1) Educ.Psy.	(2) Lit.Lang.	(3) Health	(4) Pol.Soc.
Z_m				
Credit req. (st.): EdPsy	4.587***	-0.610***	0.610	-0.581***
	(0.081)	(0.137)	(0.573)	(0.094)
Credit req. (st.): PlSc	-1.547***	-0.731***	-0.842***	-0.512***
	(0.021)	(0.162)	(0.210)	(0.111)
log(distance)	-0.016***	-0.081***	-0.055	-0.031
	(0.004)	(0.028)	(0.035)	(0.019)
X				
HS: grade (st.)	0.365***	0.194***	0.018	0.312***
	(0.011)	(0.070)	(0.102)	(0.048)
HS: humanities	0.786***	0.878***	0.362	0.574***
	(0.032)	(0.182)	(0.298)	(0.136)
HS: science	0.670***	-0.055	-0.173	0.313***
	(0.023)	(0.170)	(0.237)	(0.107)
Gender (1=female)	-0.446***	-0.454**	-0.661**	-0.548***
	(0.032)	(0.199)	(0.270)	(0.140)
Parents: graduate	0.371***	0.344*	-0.123	0.417***
	(0.029)	(0.186)	(0.315)	(0.123)
Parents: high-rank occ.	0.105***	-0.168	-0.711*	0.185
	(0.029)	(0.203)	(0.367)	(0.128)
Additional Controls		Yes		
Θ		Yes		
Constant	3.314***	-4.234***	-5.255***	-2.900***
	(0.134)	(0.702)	(1.131)	(0.457)
Observations	62,741	62,741	62,741	62,741
Pseudo R2	0.223	0.223	0.223	0.223

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Excluded category: no master.

Table B.9: Probability of choosing a master degree given a bachelor in Law

VARIABLES	(1) Econ.Mgmt.	(2) Educ.Psy.	(3) Law	(4) Pol.Soc.
Z_m				
Credit req. (st.): PlSc	-3.800***	-2.129***	-5.008***	-2.087***
	(0.149)	(0.418)	(0.154)	(0.160)
log(distance)	-0.040***	0.032	0.012	-0.032*
	(0.013)	(0.057)	(0.010)	(0.017)
X				
HS: grade (st.)	0.540***	0.233**	0.270***	0.311***
	(0.034)	(0.098)	(0.033)	(0.037)
HS: humanities	0.166	1.392***	0.814***	0.893***
	(0.115)	(0.240)	(0.097)	(0.104)
HS: science	0.796***	0.755***	0.571***	0.592***
	(0.075)	(0.227)	(0.076)	(0.084)
Gender (1=female)	-0.137	0.078	-0.640***	-0.509***
	(0.088)	(0.257)	(0.088)	(0.095)
Parents: graduate	0.592***	0.669***	0.335***	0.404***
	(0.091)	(0.242)	(0.096)	(0.102)
Parents: high-rank occ.	0.411***	0.121	-0.042	-0.001
	(0.091)	(0.258)	(0.100)	(0.106)
Additional Controls		Yes		
Θ		Yes		
Constant	-13.281***	-28.260	-12.949***	-8.243***
	(0.567)	(374.262)	(0.494)	(0.551)
Observations	10,882	10,882	10,882	10,882
Pseudo R2	0.182	0.182	0.182	0.182

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Excluded category: no master.

Table B.10: Probability of choosing a master degree given a bachelor in Literature and Languages

VARIABLES	(1) Arc.Eng.	(2) Econ.Mgmt.	(3) Educ.Psy.	(4) Lit.Lang.	(5) Pol.Soc.	(6) Sci.Stat.
Z_m						
Credit req. (st.): EdPs	3.252***	1.105***	-1.907***	-1.582***	0.219*	0.222
	(1.001)	(0.371)	(0.208)	(0.052)	(0.123)	(0.888)
Credit req. (st.): Hum	7.735***	-3.854***	0.118	-1.500***	-4.206***	0.931
	(0.970)	(0.611)	(0.513)	(0.089)	(0.207)	(0.845)
Credit req. (st.): PlSc	-1.525**	2.315***	-0.244	-0.122**	1.966***	-2.063***
	(0.633)	(0.412)	(0.330)	(0.057)	(0.138)	(0.543)
Credit req. (st.): Sci	-0.176	2.466***	-0.433	-2.117***	0.411**	-0.963
	(1.986)	(0.521)	(0.520)	(0.089)	(0.173)	(0.900)
log(distance)	-0.041	0.045**	-0.005	-0.008**	-0.030***	-0.037
	(0.042)	(0.022)	(0.015)	(0.004)	(0.006)	(0.038)
X						
HS: grade (st.)	0.580***	0.428***	-0.123***	0.565***	0.381***	0.251***
	(0.096)	(0.042)	(0.044)	(0.008)	(0.015)	(0.083)
HS: humanities	0.315	0.839***	0.367***	1.115***	0.846***	0.593***
	(0.288)	(0.108)	(0.115)	(0.021)	(0.040)	(0.224)
HS: science	0.615***	0.784***	0.250**	0.753***	0.836***	0.605***
	(0.213)	(0.093)	(0.102)	(0.019)	(0.035)	(0.187)
Gender (1=female)	-0.833***	-0.068	0.073	-0.393***	-0.740***	-1.032***
	(0.251)	(0.125)	(0.135)	(0.024)	(0.043)	(0.252)
Parents: graduate	0.769***	0.373***	0.136	0.403***	0.244***	0.439**
	(0.213)	(0.096)	(0.111)	(0.020)	(0.036)	(0.199)
Parents: high-rank occ.	0.189	0.020	-0.209*	-0.058***	0.052	-0.271
	(0.219)	(0.102)	(0.121)	(0.021)	(0.038)	(0.233)
Additional Controls			Yes			
Θ			Yes			
Constant	-3.131**	-6.094***	-1.798***	-0.913***	0.201	-7.927***
	(1.546)	(0.588)	(0.528)	(0.105)	(0.189)	(1.456)
Observations	90,681	90,681	90,681	90,681	90,681	90,681
Pseudo R2	0.108	0.108	0.108	0.108	0.108	0.108

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Excluded category: no master.

Table B.11: Probability of choosing a master degree given a bachelor in Health

VARIABLES	(1) AVGB	(2) Educ.Psy.	(3) Health
Z_m			
Credit req. (st.): AVGB	-4.685***	9.015***	-4.584***
	(0.189)	(1.686)	(0.160)
log(distance)	-0.017	-0.033	0.011*
	(0.019)	(0.023)	(0.006)
X			
HS: grade (st.)	0.403***	-0.032	0.115***
	(0.054)	(0.062)	(0.015)
HS: humanities	0.205	-0.376*	0.147***
	(0.180)	(0.219)	(0.049)
HS: science	0.401***	-0.402***	-0.193***
	(0.109)	(0.132)	(0.032)
Gender (1=female)	-0.258*	-0.006	-0.027
	(0.142)	(0.164)	(0.040)
Parents: graduate	0.186	0.309*	-0.098**
	(0.143)	(0.174)	(0.047)
Parents: high-rank occ.	0.063	0.200	-0.070
	(0.157)	(0.183)	(0.050)
Additional Controls		Yes	
Θ		Yes	
Constant	-5.587***	-1.271	-4.954***
	(0.453)	(0.806)	(0.158)
Observations	81,883	81,883	81,883
Pseudo R2	0.0591	0.0591	0.0591

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Excluded category: no master.

Table B.12: Probability of choosing a master degree given a bachelor in Political and Social Sciences

VARIABLES	(1) Econ.Mgmt.	(2) Educ.Psy.	(3) Law	(4) Lit.Lang.	(5) Pol.Soc.	(6) Sci.Stat.
Z_m						
Credit req. (st.): EcMg	-1.417**	4.024***	8.987***	1.368***	3.404***	4.183**
	(0.661)	(0.557)	(1.241)	(0.453)	(0.154)	(1.859)
Credit req. (st.): EdPs	-2.027***	-4.539***	-7.398***	-3.631***	-2.605***	-3.432***
	(0.733)	(0.345)	(0.867)	(0.681)	(0.086)	(1.207)
Credit req. (st.): Law	-1.630***	-3.461***	-7.974***	-2.296***	-2.837***	-3.141**
	(0.613)	(0.355)	(0.829)	(0.489)	(0.103)	(1.262)
Credit req. (st.): Hum	13.178***	-15.963***	-15.441***	-4.140**	-3.580***	-10.459
	(2.992)	(1.868)	(4.253)	(1.981)	(0.608)	(7.022)
Credit req. (st.): PlSc	6.158***	-8.130***	-7.197***	2.305***	-3.379***	-4.010
	(1.142)	(0.674)	(1.688)	(0.725)	(0.233)	(2.711)
log(distance)	-0.024**	-0.000	0.029***	-0.013	-0.003	0.104
	(0.011)	(0.018)	(0.010)	(0.009)	(0.004)	(0.066)
X						
HS: grade (st.)	0.433***	0.385***	0.228***	0.604***	0.468***	0.795***
	(0.029)	(0.049)	(0.031)	(0.026)	(0.010)	(0.104)
HS: humanities	0.342***	0.615***	0.053	0.864***	0.790***	-0.079
	(0.086)	(0.134)	(0.089)	(0.066)	(0.028)	(0.369)
HS: science	0.794***	0.646***	0.249***	0.483***	0.668***	0.610***
	(0.061)	(0.108)	(0.067)	(0.058)	(0.022)	(0.210)
Gender (1=female)	-0.192***	0.616***	-0.497***	-0.135**	-0.165***	-1.282***
	(0.073)	(0.145)	(0.083)	(0.067)	(0.027)	(0.266)
Parents: graduate	0.222***	0.398***	-0.015	0.467***	0.274***	0.210
	(0.068)	(0.118)	(0.082)	(0.059)	(0.025)	(0.247)
Parents: high-rank occ.	-0.070	0.039	-0.240***	-0.079	-0.078***	-0.260
	(0.072)	(0.124)	(0.089)	(0.063)	(0.026)	(0.267)
Additional Controls			Yes			
Θ			Yes			
Constant	20.289***	-29.793***	-11.597**	5.244**	-4.120***	-14.040*
	(3.374)	(1.974)	(4.706)	(2.182)	(0.667)	(7.746)
Observations	65,798	65,798	65,798	65,798	65,798	65,798
Pseudo R2	0.162	0.162	0.162	0.162	0.162	0.162

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Excluded category: no master.

Table B.13: Probability of choosing a master degree given a bachelor in Science and Statistics

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	AVGB	Arc.Eng.	Chem.Phar.	Econ.Mgmt.	Pol.Soc.	Sci.Stat.
Z_m						
Credit req. (st.): ChPh	-14.708***	2.612***	-81.978	7.585***	1.568*	-4.359***
	(1.878)	(0.842)	(2,381.101)	(1.886)	(0.911)	(0.173)
Credit req. (st.): Sci	-0.320	1.618***	-21.711	0.989***	0.567*	-0.879***
	(1.022)	(0.272)	(709.538)	(0.279)	(0.335)	(0.062)
log(distance)	-0.042*	-0.021	-0.067	0.072	-0.006	-0.005
	(0.024)	(0.049)	(0.042)	(0.076)	(0.037)	(0.009)
X						
HS: grade (st.)	0.031	0.433***	0.386***	0.606***	0.183**	0.736***
	(0.052)	(0.102)	(0.090)	(0.100)	(0.086)	(0.019)
HS: humanities	1.326***	0.698*	0.271	0.928**	0.736**	0.829***
	(0.187)	(0.363)	(0.339)	(0.432)	(0.364)	(0.084)
HS: science	1.020***	-0.047	0.370**	1.046***	0.377**	0.955***
	(0.117)	(0.219)	(0.186)	(0.205)	(0.179)	(0.038)
Gender (1=female)	1.020***	-0.312	0.484*	0.197	-0.662***	0.140***
	(0.143)	(0.276)	(0.252)	(0.259)	(0.233)	(0.053)
Parents: graduate	-0.181	0.530**	0.316	0.140	0.110	0.365***
	(0.119)	(0.235)	(0.198)	(0.232)	(0.216)	(0.045)
Parents: high-rank occ.	0.021	0.610**	-0.110	0.252	0.019	0.008
	(0.131)	(0.239)	(0.227)	(0.243)	(0.231)	(0.050)
Additional Controls			Yes			
Θ			Yes			
Constant	9.007***	-4.773***	32.378	-13.920***	-1.120	2.247***
	(0.791)	(1.286)	(896.781)	(2.239)	(0.994)	(0.222)
Observations	20,721	20,721	20,721	20,721	20,721	20,721
Pseudo R2	0.300	0.300	0.300	0.300	0.300	0.300

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Excluded category: no master.

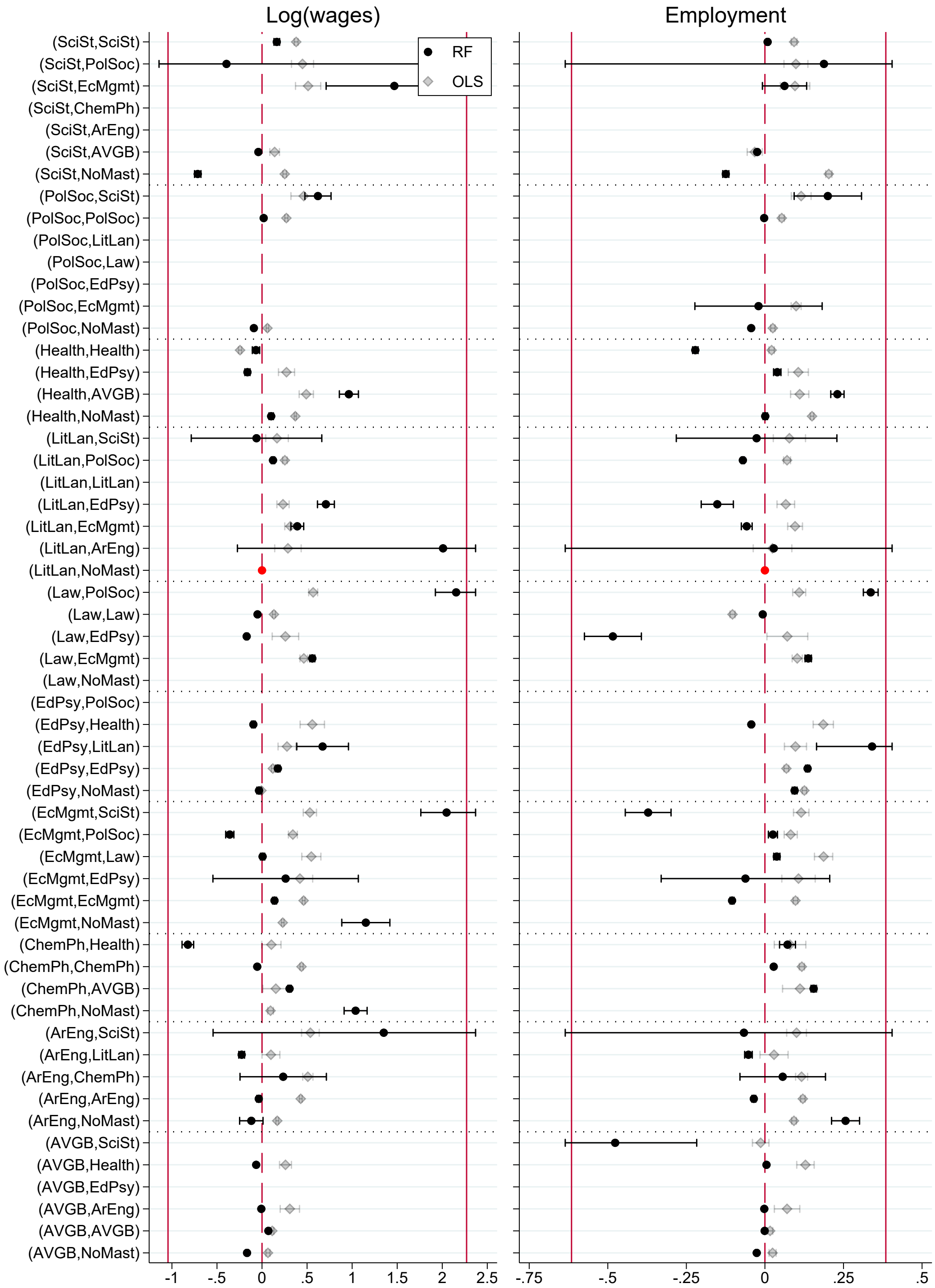
B.4 Additional Results

Table B.14: Summary of treatments D_{jm} and probabilities P_{jm}

#	(j, m)	D_{jm}		P_{jm}			$P_{jm} - D_{jm}$
		Mean	Std. Dev.	Mean	Std. Dev.	Max	
1	(AVGB, No Master)	0.0128	(0.1124)	0.0019	(0.0057)	0.1613	-0.0109
2	(AVGB, AVGB)	0.0401	(0.1963)	0.0141	(0.023)	0.1944	-0.026
3	(AVGB, Arch.Eng.)	0.0003	(0.0183)	0	(0.0002)	0.0236	-0.0003
4	(AVGB, Educ.Psy.)	0.0003	(0.0166)	0.0071	(0.0205)	0.1911	0.0069
5	(AVGB, Pol.Soc.)	0.0009	(0.0308)	0.0001	(0.0008)	0.0586	-0.0009
6	(AVGB, Sci.Stat.)	0.0014	(0.0377)	0.0326	(0.0317)	0.2041	0.0312
7	(Arch.Eng., No Master)	0.034	(0.1812)	0.0631	(0.0722)	0.5837	0.0291
8	(Arch.Eng., Arch.Eng.)	0.1217	(0.327)	0.0939	(0.1237)	0.7117	-0.0278
9	(Arch.Eng., Chem.Pharm.)	0.0012	(0.0344)	0.0004	(0.0023)	0.066	-0.0008
10	(Arch.Eng., Lit.Lang.)	0.0004	(0.0209)	0.0001	(0.0009)	0.1028	-0.0004
11	(Arch.Eng., Sci.Stat.)	0.0004	(0.0196)	0.0002	(0.0012)	0.0995	-0.0002
12	(Chem.Pharm., No Master)	0.0059	(0.0769)	0.0264	(0.023)	0.2088	0.0204
13	(Chem.Pharm., AVGB)	0.0002	(0.0134)	0.0014	(0.0031)	0.0839	0.0012
14	(Chem.Pharm., Chem.Pharm.)	0.0315	(0.1746)	0.0085	(0.0216)	0.1834	-0.023
15	(Chem.Pharm., Health)	0.0004	(0.0199)	0.0017	(0.0057)	0.1438	0.0013
16	(Econ.Mgmt., No Master)	0.0424	(0.2015)	0.0724	(0.068)	0.5057	0.03
17	(Econ.Mgmt., Econ.Mgmt.)	0.0705	(0.256)	0.042	(0.0611)	0.462	-0.0285
18	(Econ.Mgmt., Educ.Psy.)	0.0002	(0.0137)	0.0001	(0.0002)	0.0115	-0.0001
19	(Econ.Mgmt., Law)	0.0003	(0.0178)	0.0001	(0.0006)	0.0641	-0.0002
20	(Econ.Mgmt., Pol.Soc.)	0.0018	(0.0419)	0.0011	(0.002)	0.0523	-0.0007
21	(Econ.Mgmt., Sci.Stat.)	0.0007	(0.0264)	0.0002	(0.0006)	0.0281	-0.0005
22	(Educ.Psy., No Master)	0.0435	(0.204)	0.0449	(0.0559)	0.5477	0.0014
23	(Educ.Psy., Educ.Psy.)	0.0703	(0.2556)	0.0547	(0.061)	0.5451	-0.0156
24	(Educ.Psy., Lit.Lang.)	0.0004	(0.0195)	0.0035	(0.0116)	0.313	0.0031
25	(Educ.Psy., Health)	0.0002	(0.0138)	0.0002	(0.0008)	0.0644	0
26	(Educ.Psy., Pol.Soc.)	0.0008	(0.0286)	0.0119	(0.0359)	0.4998	0.0111
27	(Law, No Master)	0.0123	(0.1101)	0.052	(0.0512)	0.5102	0.0398
28	(Law, Econ.Mgmt.)	0.0022	(0.0472)	0.0097	(0.0168)	0.3091	0.0074
29	(Law, Educ.Psy.)	0.0002	(0.0139)	0.0009	(0.0025)	0.0933	0.0007
30	(Law, Law)	0.0713	(0.2573)	0.0186	(0.0435)	0.4688	-0.0527
31	(Law, Pol.Soc.)	0.0017	(0.0409)	0.0065	(0.0121)	0.2348	0.0048
32	(Lit.Lang., No Master)	0.0585	(0.2346)	0.0649	(0.0539)	0.5661	0.0064
33	(Lit.Lang., Arch.Eng.)	0.0002	(0.0136)	0.0037	(0.0145)	0.5213	0.0035
34	(Lit.Lang., Econ.Mgmt.)	0.0011	(0.0325)	0.0007	(0.0019)	0.0706	-0.0004
35	(Lit.Lang., Educ.Psy.)	0.0009	(0.0301)	0.0027	(0.0058)	0.1723	0.0018
36	(Lit.Lang., Lit.Lang.)	0.0686	(0.2527)	0.0614	(0.0604)	0.6429	-0.0071
37	(Lit.Lang., Pol.Soc.)	0.0088	(0.0935)	0.0046	(0.0071)	0.1531	-0.0042
38	(Lit.Lang., Sci.Stat.)	0.0003	(0.0159)	0.0002	(0.0016)	0.1725	0
39	(Health, No Master)	0.1155	(0.3196)	0.1337	(0.1242)	0.7477	0.0183
40	(Health, AVGB)	0.0006	(0.0248)	0.0025	(0.0073)	0.3174	0.0018
41	(Health, Educ.Psy.)	0.0005	(0.0218)	0.0004	(0.0011)	0.0337	0
42	(Health, Health)	0.0428	(0.2024)	0.0227	(0.041)	0.6067	-0.0201
43	(Pol.Soc., No Master)	0.0534	(0.2248)	0.0124	(0.0309)	0.3077	-0.041
44	(Pol.Soc., Econ.Mgmt.)	0.0024	(0.0487)	0.0596	(0.0545)	0.3717	0.0572
45	(Pol.Soc., Educ.Psy.)	0.0009	(0.0302)	0.008	(0.0259)	0.2688	0.0071
46	(Pol.Soc., Law)	0.002	(0.0452)	0.0104	(0.0276)	0.3019	0.0083
47	(Pol.Soc., Lit.Lang.)	0.003	(0.0544)	0.0044	(0.0141)	0.2164	0.0014
48	(Pol.Soc., Pol.Soc.)	0.0386	(0.1927)	0.0054	(0.0155)	0.2249	-0.0332
49	(Pol.Soc., Sci.Stat.)	0.0002	(0.0131)	0.0003	(0.0015)	0.0539	0.0001
50	(Sci.Stat., No Master)	0.0131	(0.1137)	0.0132	(0.0174)	0.2344	0.0001
51	(Sci.Stat., AVGB)	0.0015	(0.0393)	0.0001	(0.0018)	0.0881	-0.0014
52	(Sci.Stat., Arch.Eng.)	0.0002	(0.0132)	0.0011	(0.0026)	0.0651	0.0009
53	(Sci.Stat., Chem.Pharm.)	0.0003	(0.0167)	0.0097	(0.0217)	0.2926	0.0094
54	(Sci.Stat., Econ.Mgmt.)	0.0002	(0.0137)	0.0006	(0.0014)	0.0462	0.0004
55	(Sci.Stat., Pol.Soc.)	0.0002	(0.0156)	0.0005	(0.0011)	0.055	0.0002
56	(Sci.Stat., Sci.Stat.)	0.0161	(0.1257)	0.0064	(0.0115)	0.218	-0.0096

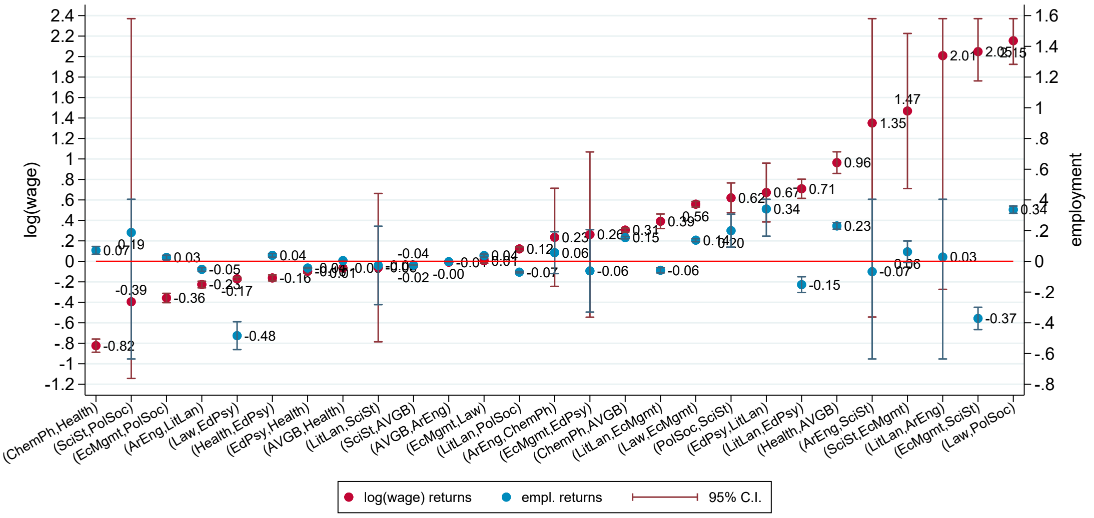
Summary statistics for the vector of treatments D_{jm} and probabilities P_{jm} for 56 combinations of bachelor's and master's degrees. Treatments D_{jm} take values 0 and 1. The minimum value for instruments P_{jm} is, hence the omission. Sums calculated on 655 847 observations.

Figure B.2: Comparison of OLS coefficients γ and reduced form treatment effects



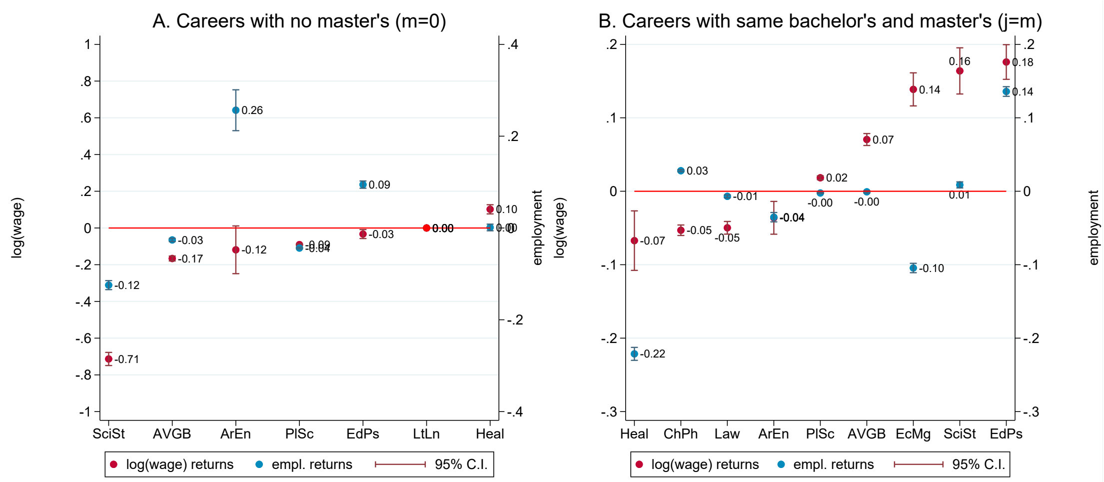
Black markers indicate reduced form (RF) results (2.4), grey markers indicate OLS results (2.5). Whiskers denote 95% CIs and the red dot denotes the baseline (Lit.Lang., No Master). Red lines denote credible boundaries for the treatment effects. $TE(\ln(\text{wage})) \in [-1.04, 2.27]$, employment: $TE(\text{employment}) \in [-0.62, 0.38]$.

Figure B.3: Comparison of log wage and employment returns for all multidisciplinary careers



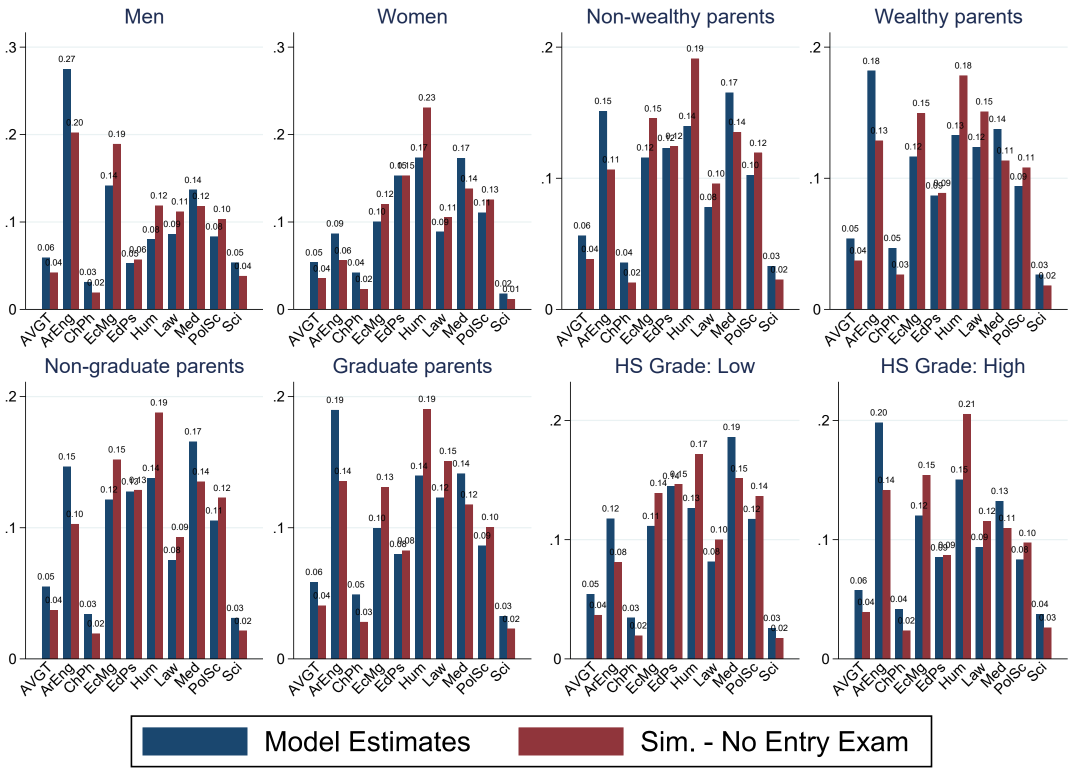
Comparison of log wage returns (red, left vertical axis) and returns to employment (blue, right vertical axis). Axes are centered around 0. The excluded category is (Lit.Lang., No Master) centered at 0. Any missing returns could not be estimated for both outcomes. Panel B presents returns to specialized careers with the same bachelor's and master's. The order follows the ranking of log wage returns from lowest to highest.

Figure B.4: Comparison of log wage and employment returns for non-multidisciplinary careers



Comparison of log wage returns (red, left vertical axis) and returns to employment (blue, right vertical axis). Axes are centered around 0. Panel A presents labor market returns for careers with no master, where (Lit.Lang., No Master) denotes the excluded category centered at 0. Any missing returns could not be estimated for both outcomes. Panel B presents returns to specialized careers with the same bachelor's and masters. In both panels, the order follows the ranking of log wage returns from lowest to highest.

Figure B.5: Simulation 1 – Decomposition by individual characteristics



B.5 Methodological notes

B.5.1 Universities

The universities that I consider are the following: Politecnico di Ancona, Bari, Politecnico di Bari, Basilicata, Bergamo, Bologna, Bolzano, Brescia, Cagliari, Calabria, Camerino, Campania - Luigi Vanvitelli, Cassino e Lazio Meridionale, Catania, Catanzaro, Chieti e Pescara, Enna Kore, Ferrara, Firenze, Foggia, Genova, Insubira, L'Aquila, LIUC Castellanza, Macerata, Messina, Milano Bicocca, Milano IULM, Milano Statale, Milano Vita-Salute S. Raffaele, Modena e Reggio Emilia, Molise, Napoli - Federico II, Napoli - Seconda Università, Napoli - L'Orientale, Napoli - Parthenope, Padova, Palermo, Parma, Pavia, Perugia, Università per Stranieri di Perugia, Piemonte Orientale, Pisa, Reggio Calabria Mediterranea, Roma - Campus Bio-Medico, Roma LUMSA, Roma Foro Italico, Roma Tre, Roma - La Sapienza, Roma - Tor Vergata, Salento, Salerno, Sannio e Benevento, Sassari, Siena, Università per Stranieri di Siena, Teramo, Torino, Politecnico di Torino, Trento, Trieste, Udine, Urbino, Viterbo Tuscia, Valle D'Aosta Venezia - Ca' Foscari, Venezia - IUAV, Verona.

Some universities which are not in this list may nonetheless appear in the dataset (e.g. Milano Bocconi). The reason is that students appear in the dataset if they graduated (master) from a university in the consortium, yet information is collected also for their bachelor which may differ. Only about 5% of students in the sample switches institution throughout their career.

B.5.2 Degrees and groups

Here, I present the exact pooling of degrees into groups. The allocation has been done by the AlmaLaurea consortium. A few groups of degrees were further grouped to improve estimation: agriculture and veterinary was grouped with geology and biology, architecture with engineering, teaching with physical education and psychology, and literature with languages. Information on an additional group – defense and security – was dropped as access into these degrees is managed differently from standard university degrees.

Table B.15: Degree grouping

Code	**Description**
1. Agriculture, veterinarian sciences, geology, biology	
L-2	BIOTECNOLOGIE
L-13	SCIENZE BIOLOGICHE
L-25	SCIENZE E TECNOLOGIE AGRARIE E FORESTALI
L-26	SCIENZE E TECNOLOGIE AGRO-ALIMENTARI
L-32	SCIENZE E TECNOLOGIE PER L'AMBIENTE E LA NATURA
L-34	SCIENZE GEOLOGICHE
L-38	SCIENZE ZOOTECNICHE E TECNOLOGIE DELLE PRODUZIONI ANIMALI
LM-6	BIOLOGIA
LM-7	BIOTECNOLOGIE AGRARIE
LM-8	BIOTECNOLOGIE INDUSTRIALI
LM-9	BIOTECNOLOGIE MEDICHE, VETERINARIE E FARMACEUTICHE
LM-42	MEDICINA VETERINARIA
LM-69	SCIENZE E TECNOLOGIE AGRARIE
LM-70	SCIENZE E TECNOLOGIE ALIMENTARI
LM-73	SCIENZE E TECNOLOGIE FORESTALI ED AMBIENTALI
LM-74	SCIENZE E TECNOLOGIE GEOLOGICHE
LM-75	SCIENZE E TECNOLOGIE PER L'AMBIENTE E IL TERRITORIO

Code	Description
LM-86	SCIENZE ZOOTECNICHE E TECNOLOGIE ANIMALI
2. Architecture and Engineering	
L-4	DISEGNO INDUSTRIALE
L-7	INGEGNERIA CIVILE E AMBIENTALE
L-8	INGEGNERIA DELL'INFORMAZIONE
L-9	INGEGNERIA INDUSTRIALE
L-17	SCIENZE DELL'ARCHITETTURA
L-21	SCIENZE DELLA PIANIFICAZIONE TERRITORIALE, URBANISTICA, PAESAGGISTICA E AMBIENTALE
L-23	SCIENZE E TECNICHE DELL'EDILIZIA
LM-3	ARCHITETTURA DEL PAESAGGIO
LM-4	ARCHITETTURA E INGEGNERIA EDILE-ARCHITETTURA
LM-12	DESIGN
LM-20	INGEGNERIA AEROSPAZIALE E ASTRONAUTICA
LM-21	INGEGNERIA BIOMEDICA
LM-22	INGEGNERIA CHIMICA
LM-23	INGEGNERIA CIVILE
LM-24	INGEGNERIA DEI SISTEMI EDILIZI
LM-25	INGEGNERIA DELL'AUTOMAZIONE
LM-26	INGEGNERIA DELLA SICUREZZA
LM-27	INGEGNERIA DELLE TELECOMUNICAZIONI
LM-28	INGEGNERIA ELETTRICA
LM-29	INGEGNERIA ELETTRONICA
LM-30	INGEGNERIA ENERGETICA E NUCLEARE
LM-31	INGEGNERIA GESTIONALE
LM-32	INGEGNERIA INFORMATICA
LM-33	INGEGNERIA MECCANICA
LM-34	INGEGNERIA NAVALE
LM-35	INGEGNERIA PER L'AMBIENTE E IL TERRITORIO
LM-44	MODELLISTICA MATEMATICO-FISICA PER L'INGEGNERIA
LM-48	PIANIFICAZIONE TERRITORIALE URBANISTICA E AMBIENTALE
LM-53	SCIENZA E INGEGNERIA DEI MATERIALI
3. Chemistry and Pharmacy	
L-27	SCIENZE E TECNOLOGIE CHIMICHE
L-29	SCIENZE E TECNOLOGIE FARMACEUTICHE
LM-13	FARMACIA E FARMACIA INDUSTRIALE
LM-54	SCIENZE CHIMICHE
LM-71	SCIENZE E TECNOLOGIE DELLA CHIMICA INDUSTRIALE

Code	Description
	4. Economics and Management
L-15	SCIENZE DEL TURISMO
L-16	SCIENZE DELL'AMMINISTRAZIONE E DELL'ORGANIZZAZIONE
L-18	SCIENZE DELL'ECONOMIA E DELLA GESTIONE AZIENDALE
L-33	SCIENZE ECONOMICHE
LM-16	FINANZA
LM-56	SCIENZE DELL'ECONOMIA
LM-76	SCIENZE ECONOMICHE PER L'AMBIENTE E LA CULTURA
LM-77	SCIENZE ECONOMICO-AZIENDALI
	5. Teaching, Physical Education and Psychology
L-19	SCIENZE DELL'EDUCAZIONE E DELLA FORMAZIONE
L-22	SCIENZE DELLE ATTIVITA MOTORIE E SPORTIVE
L-24	SCIENZE E TECNICHE PSICOLOGICHE
LM-47	ORGANIZZAZIONE E GESTIONE DEI SERVIZI PER LO SPORT E LE ATTIVITA MOTORIE
LM-50	PROGRAMMAZIONE E GESTIONE DEI SERVIZI EDUCATIVI
LM-51	PSICOLOGIA
LM-55	SCIENZE COGNITIVE
LM-57	SCIENZE DELL'EDUCAZIONE DEGLI ADULTI E DELLA FORMAZIONE CONTINUA
LM-67	SCIENZE E TECNICHE DELLE ATTIVITA MOTORIE PREVENTIVE E ADATTATE
LM-68	SCIENZE E TECNICHE DELLO SPORT
LM-85	SCIENZE PEDAGOGICHE
LM-93	TEORIE E METODOLOGIE DELL'E-LEARNING E DELLA MEDIA EDUCATION
	6. Law
L-14	SCIENZE DEI SERVIZI GIURIDICI
LMG-1	GIURISPRUDENZA
	7. Literature and Languages
L-1	BENI CULTURALI
L-3	DISCIPLINE DELLE ARTI FIGURATIVE, DELLA MUSICA, DELLO SPETTACOLO E DELLA MODA (DAMS)
L-5	FILOSOFIA

Code	Description
L-6	GEOGRAFIA
L-10	LETTERE
L-11	LINGUE E CULTURE MODERNE
L-12	MEDIAZIONE LINGUISTICA
L-42	STORIA
L-43	TECNOLOGIE PER LA CONSERVAZIONE E IL RESTAURO DEI BENI CULTURALI
LM-1	ANTROPOLOGIA CULTURALE ED ETNOLOGIA
LM-2	ARCHEOLOGIA
LM-5	ARCHIVISTICA E BIBLIOTECONOMIA
LM-10	CONSERVAZIONE DEI BENI ARCHITETTONICI E AMBIENTALI
LM-11	CONSERVAZIONE E RESTAURO DEI BENI CULTURALI
LM-14	FILOLOGIA MODERNA
LM-15	FILOLOGIA, LETTERATURE E STORIA DELL'ANTICHITA
LM-36	LINGUE E LETTERATURE DELL'AFRICA E DELL'ASIA
LM-37	LINGUE E LETTERATURE MODERNE EUROPEE E AMERICANE
LM-38	LINGUE MODERNE PER LA COMUNICAZIONE E LA COOPERAZIONE
LM-39	LINGUISTICA
LM-45	MUSICOLOGIA E BENI MUSICALI
LM-65	SCIENZE DELLO SPETTACOLO E PRODUZIONE MULTIMEDIALE
LM-78	SCIENZE FILOSOFICHE
LM-80	SCIENZE GEOGRAFICHE
LM-84	SCIENZE STORICHE
LM-89	STORIA DELL'ARTE
LM-94	TRADUZIONE SPECIALISTICA E INTERPRETARIATO
8. Health and Medicine	
L/SNT-1	SCIENZE INFERMIERISTICHE E OSTETRICHE
L/SNT-2	SCIENZE RIABILITATIVE DELLE PROFESSIONI SANITARIE
L/SNT-3	SCIENZE DELLE PROFESSIONI SANITARIE TECNICHE
L/SNT-4	SCIENZE DELLE PROFESSIONI SANITARIE DELLA PREVENZIONE
LM/SNT-1	SCIENZE INFERMIERISTICHE E OSTETRICHE
LM/SNT-2	SCIENZE RIABILITATIVE DELLE PROFESSIONI SANITARIE
LM/SNT-3	SCIENZE DELLE PROFESSIONI SANITARIE TECNICHE

Code	Description
LM/SNT-4	SCIENZE DELLE PROFESSIONI SANITARIE DELLA PREVENZIONE
LM-41	MEDICINA E CHIRURGIA
LM-46	ODONTOIATRIA E PROTESI DENTARIA
LM-61	SCIENZE DELLA NUTRIZIONE UMANA
	9. Political and social sciences
L-20	SCIENZE DELLA COMUNICAZIONE
L-36	SCIENZE POLITICHE E DELLE RELAZIONI INTERNAZIONALI
L-37	SCIENZE SOCIALI PER LA COOPERAZIONE, LO SVILUPPO E LA PACE
L-39	SERVIZIO SOCIALE
L-40	SOCIOLOGIA
LM-19	INFORMAZIONE E SISTEMI EDITORIALI
LM-49	PROGETTAZIONE E GESTIONE DEI SISTEMI TURISTICI
LM-52	RELAZIONI INTERNAZIONALI
LM-59	SCIENZE DELLA COMUNICAZIONE PUBBLICA, D'IMPRESA E PUBBLICITÀ
LM-62	SCIENZE DELLA POLITICA
LM-63	SCIENZE DELLE PUBBLICHE AMMINISTRAZIONI
LM-81	SCIENZE PER LA COOPERAZIONE ALLO SVILUPPO
LM-87	SERVIZIO SOCIALE E POLITICHE SOCIALI
LM-88	SOCIOLOGIA E RICERCA SOCIALE
LM-90	STUDI EUROPEI
LM-91	TECNICHE E METODI PER LA SOCIETA DELL'INFORMAZIONE
LM-92	TEORIE DELLA COMUNICAZIONE
	10. Science and Statistics
L-28	SCIENZE E TECNOLOGIE DELLA NAVIGAZIONE
L-30	SCIENZE E TECNOLOGIE FISICHE
L-31	SCIENZE E TECNOLOGIE INFORMATICHE
L-35	SCIENZE MATEMATICHE
LM-17	FISICA
LM-18	INFORMATICA
LM-40	MATEMATICA
L-41	STATISTICA
LM-43	METODOLOGIE INFORMATICHE PER LE DISCIPLINE UMANISTICHE
LM-58	SCIENZE DELL'UNIVERSO
LM-60	SCIENZE DELLA NATURA

Code	**Description**
LM-66	SICUREZZA INFORMATICA
LM-72	SCIENZE E TECNOLOGIE DELLA NAVIGAZIONE
LM-82	SCIENZE STATISTICHE
LM-83	SCIENZE STATISTICHE ATTUARIALI E FINANZIARIE

Note: Prefix L- refers to bachelor degrees, LM- to master degrees.

APPENDIX C

Appendix to Chapter 3: Additional Descriptives and Results, Notes on Gerrymandering

C.1 Additional Results

Figure C1: Population shares in split ZCTAs

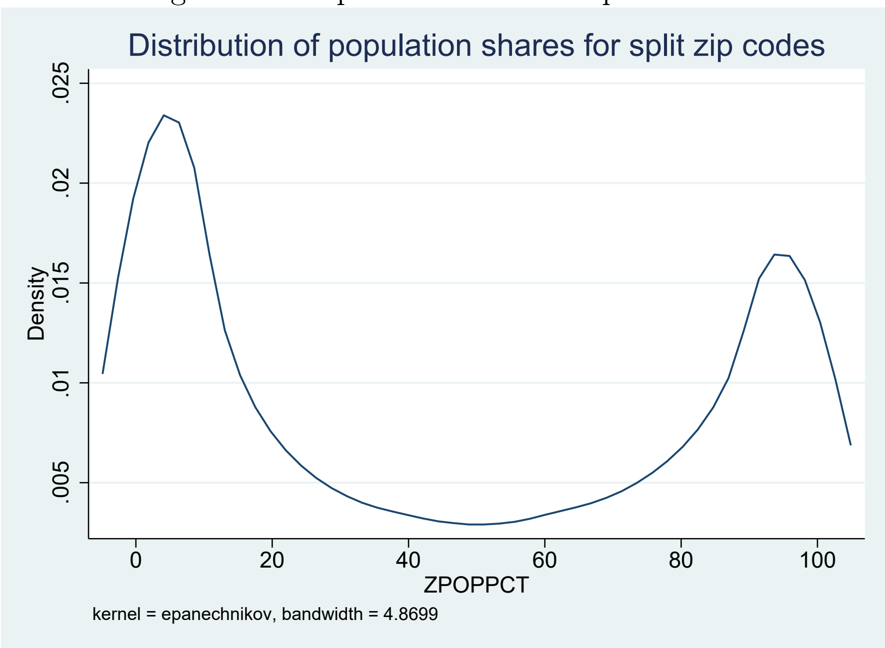


Table C1: Ideological position (DW-NOMINATE scores) for all members of 109th-115th Congresses by party

Congress	Mean	St. Dev.	Min.	Max.	p(25)	p(50)	p(75)	Obs.
Republican Representatives								
109	0.421	0.143	0.103	0.863	0.316	0.408	0.523	232
110	0.436	0.143	0.129	0.863	0.333	0.426	0.538	202
111	0.456	0.145	0.133	0.913	0.348	0.441	0.555	178
112	0.469	0.147	0.164	0.913	0.352	0.470	0.577	242
113	0.482	0.147	0.164	0.913	0.367	0.491	0.591	234
114	0.480	0.149	0.164	0.829	0.362	0.490	0.600	247
115	0.487	0.149	0.164	0.931	0.374	0.502	0.600	241
All	0.463	0.148	0.103	0.931	0.349	0.456	0.576	1576
Democratic Representatives								
109	-0.387	0.124	-0.683	-0.045	-0.473	-0.390	-0.302	202
110	-0.369	0.132	-0.683	-0.045	-0.463	-0.379	-0.282	233
111	-0.349	0.145	-0.683	0.088	-0.448	-0.350	-0.262	257
112	-0.395	0.123	-0.683	-0.070	-0.478	-0.401	-0.320	193
113	-0.383	0.115	-0.683	-0.104	-0.460	-0.390	-0.306	201
114	-0.395	0.108	-0.683	-0.104	-0.463	-0.398	-0.322	188
115	-0.390	0.113	-0.692	-0.104	-0.460	-0.396	-0.310	194
All	-0.379	0.126	-0.692	0.088	-0.464	-0.389	-0.301	1468

Table C2: Shifts in Democrat vote margin in US House general elections with above/below threshold shares.

VARIABLES	USH (D) (1)	USH (D) (2)	USH (D) (3)	USH (D) (4)	USH (D) (5)
Share Close Thresh - Above (5)	-0.235				
	(0.152)				
Share Close Thresh - Below (5)	-0.461				
	(0.315)				
Share Close Thresh - Above (10)		-0.263*			
		(0.148)			
Share Close Thresh - Below (10)		-0.246			
		(0.183)			
Share Close Thresh - Above (15)			-0.258**		
			(0.125)		
Share Close Thresh - Below (15)			-0.272*		
			(0.156)		
Share Close Thresh - Above (20)				-0.279**	
				(0.114)	
Share Close Thresh - Below (20)				-0.169	
				(0.128)	
Share Close Thresh - Above (25)					-0.276***
					(0.105)
Share Close Thresh - Below (25)					-0.139
					(0.121)
Share China Import	0.016	0.017	0.017	0.017	0.017
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Share White	-0.693***	-0.695***	-0.698***	-0.698***	-0.700***
	(0.019)	(0.020)	(0.020)	(0.020)	(0.020)
Share Female (voting age)	0.971***	0.970***	0.967***	0.966***	0.963***
	(0.179)	(0.179)	(0.179)	(0.179)	(0.179)
Observations	12,053,931	12,053,931	12,053,931	12,053,931	12,053,931
R-squared	0.420	0.420	0.420	0.421	0.421

Clustered SE at county x congressional district
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C3: Shifts in Republican vote margin in US House general elections with above/below threshold shares.

VARIABLES	USH (R) (1)	USH (R) (2)	USH (R) (3)	USH (R) (4)	USH (R) (5)
Share Close Thresh - Above (5)	0.280*				
	(0.156)				
Share Close Thresh - Below (5)	0.445				
	(0.308)				
Share Close Thresh - Above (10)		0.257*			
		(0.148)			
Share Close Thresh - Below (10)		0.266			
		(0.185)			
Share Close Thresh - Above (15)			0.248**		
			(0.126)		
Share Close Thresh - Below (15)			0.292*		
			(0.158)		
Share Close Thresh - Above (20)				0.262**	
				(0.114)	
Share Close Thresh - Below (20)				0.186	
				(0.131)	
Share Close Thresh - Above (25)					0.260**
					(0.104)
Share Close Thresh - Below (25)					0.156
					(0.124)
Share China Import	-0.018	-0.018	-0.019	-0.019	-0.019
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Share White	0.700***	0.702***	0.705***	0.705***	0.707***
	(0.019)	(0.020)	(0.020)	(0.020)	(0.020)
Share Female (voting age)	-0.974***	-0.972***	-0.969***	-0.969***	-0.966***
	(0.183)	(0.183)	(0.183)	(0.183)	(0.183)
Observations	12,053,931	12,053,931	12,053,931	12,053,931	12,053,931
R-squared	0.435	0.435	0.435	0.435	0.435

Clustered SE at county x congressional district
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C4: Shifts in Democrat vote margin in US House general elections with proximity to threshold shares.

VARIABLES	USH (D) (1)	USH (D) (2)	USH (D) (3)	USH (D) (4)	USH (D) (5)
Share Close Thresh (5)	-0.327**				
	(0.162)				
Share Close Thresh (10)		-0.255*			
		(0.133)			
Share Close Thresh (15)			-0.264**		
			(0.118)		
Share Close Thresh (20)				-0.227**	
				(0.101)	
Share Close Thresh (25)					-0.212**
					(0.094)
Share China Import	0.016	0.017	0.017	0.017	0.017
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Share White	-0.693***	-0.695***	-0.698***	-0.699***	-0.700***
	(0.019)	(0.020)	(0.020)	(0.020)	(0.020)
Share Female (voting age)	0.972***	0.969***	0.967***	0.967***	0.964***
	(0.179)	(0.179)	(0.179)	(0.179)	(0.179)
Observations	12,053,931	12,053,931	12,053,931	12,053,931	12,053,931
R-squared	0.420	0.420	0.420	0.420	0.421

Clustered SE at county x congressional district
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C5: Shifts in Republican vote margin in US House general elections with proximity to threshold shares.

VARIABLES	USH (R) (1)	USH (R) (2)	USH (R) (3)	USH (R) (4)	USH (R) (5)
Share Close Thresh (5)	0.347**				
	(0.165)				
Share Close Thresh (10)		0.261*			
		(0.135)			
Share Close Thresh (15)			0.268**		
			(0.119)		
Share Close Thresh (20)				0.226**	
				(0.103)	
Share Close Thresh (25)					0.212**
					(0.095)
Share China Import	-0.018	-0.018	-0.019	-0.019	-0.019
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Share White	0.700***	0.702***	0.705***	0.706***	0.707***
	(0.019)	(0.020)	(0.020)	(0.020)	(0.020)
Share Female (voting age)	-0.975***	-0.972***	-0.969***	-0.969***	-0.966***
	(0.183)	(0.183)	(0.183)	(0.183)	(0.184)
Observations	12,053,931	12,053,931	12,053,931	12,053,931	12,053,931
R-squared	0.435	0.435	0.435	0.435	0.435

Clustered SE at county x congressional district

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

Table C6: Shifts in Poole-Rosenthal DW-NOMINATE scores for winning candidates in US House general elections with above/below threshold shares.

	Republican-winning districts					Democrat-winning districts				
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Share Close Thresh - Above (5)	-0.138					0.500*				
	(0.089)					(0.259)				
Share Close Thresh - Below (5)	-0.021					0.446				
	(0.191)					(0.300)				
Share Close Thresh - Above (10)		-0.075					0.428**			
		(0.088)					(0.205)			
Share Close Thresh - Below (10)		0.051					0.451**			
		(0.141)					(0.221)			
Share Close Thresh - Above (15)			-0.134					0.478***		
			(0.082)					(0.168)		
Share Close Thresh - Below (15)			0.104					0.465**		
			(0.117)					(0.186)		
Share Close Thresh - Above (20)				-0.114					0.397***	
				(0.075)					(0.145)	
Share Close Thresh - Below (20)				0.043					0.487***	
				(0.090)					(0.170)	
Share Close Thresh - Above (25)					-0.080					0.361***
					(0.068)					(0.129)
Share Close Thresh - Below (25)					0.037					0.469***
					(0.087)					(0.156)
Share China Import	0.023***	0.022***	0.023***	0.023***	0.023***	0.015	0.014	0.014	0.013	0.013
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Share White	0.047*	0.048*	0.048*	0.047*	0.048*	0.275***	0.280***	0.286***	0.290***	0.293***
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.021)	(0.021)	(0.021)	(0.022)	(0.022)
Share Female (voting age)	0.147	0.148	0.146	0.144	0.145	-0.412*	-0.410*	-0.409*	-0.408*	-0.406*
	(0.184)	(0.184)	(0.184)	(0.184)	(0.184)	(0.238)	(0.238)	(0.237)	(0.236)	(0.236)
Observations	6,275,158	6,275,158	6,275,158	6,275,158	6,275,158	5,867,112	5,867,112	5,867,112	5,867,112	5,867,112
R-squared	0.052	0.052	0.053	0.053	0.052	0.225	0.226	0.228	0.229	0.231

Clustered SE at county x congressional district
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C7: Shifts in Poole-Rosenthal DW-NOMINATE scores for winning candidates in US House general elections with proximity to threshold shares.

	Democrat-winning candidates					Republican-winning candidates				
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Share Close Thresh (5)	0.469**					-0.104				
	(0.215)					(0.094)				
Share Close Thresh (10)		0.441***					-0.026			
		(0.159)					(0.089)			
Share Close Thresh (15)			0.471***					-0.035		
			(0.145)					(0.080)		
Share Close Thresh (20)				0.441***					-0.043	
				(0.125)					(0.068)	
Share Close Thresh (25)					0.413***					-0.027
					(0.113)					(0.064)
Share China Import	0.015	0.014	0.014	0.013	0.013	0.023***	0.023***	0.023***	0.023***	0.023***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Share White	0.275***	0.280***	0.286***	0.289***	0.293***	0.047*	0.048*	0.047*	0.047*	0.047*
	(0.021)	(0.021)	(0.021)	(0.021)	(0.022)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
Share Female (voting age)	-0.412*	-0.410*	-0.409*	-0.408*	-0.405*	0.146	0.147	0.146	0.145	0.145
	(0.238)	(0.238)	(0.237)	(0.236)	(0.236)	(0.184)	(0.184)	(0.184)	(0.184)	(0.184)
Observations	5,867,112	5,867,112	5,867,112	5,867,112	5,867,112	6,275,158	6,275,158	6,275,158	6,275,158	6,275,158
R-squared	0.225	0.226	0.228	0.229	0.230	0.052	0.052	0.052	0.052	0.052

Clustered SE at county x congressional district
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C8: Shifts in Poole-Rosenthal DW-NOMINATE scores for all candidates in US House general elections with proximity to threshold shares.

VARIABLES	(1)	(2)	(3)	(4)	(5)
Share Close Thresh (5)	0.639*				
	(0.347)				
Share Close Thresh (10)		0.577**			
		(0.287)			
Share Close Thresh (15)			0.509**		
			(0.258)		
Share Close Thresh (20)				0.406*	
				(0.225)	
Share Close Thresh (25)					0.392*
					(0.206)
Share China Import	0.031	0.030	0.030	0.029	0.029
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
Share White	1.093***	1.099***	1.103***	1.103***	1.107***
	(0.048)	(0.048)	(0.048)	(0.048)	(0.049)
Share Female (voting age)	-1.047**	-1.040**	-1.037**	-1.037**	-1.032**
	(0.434)	(0.434)	(0.434)	(0.434)	(0.434)
Observations	12,146,679	12,146,679	12,146,679	12,146,679	12,146,679
R-squared	0.296	0.296	0.296	0.296	0.296

Clustered SE at county x congressional district

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

Figure C2: Thresholds: discontinuity in total credit amount

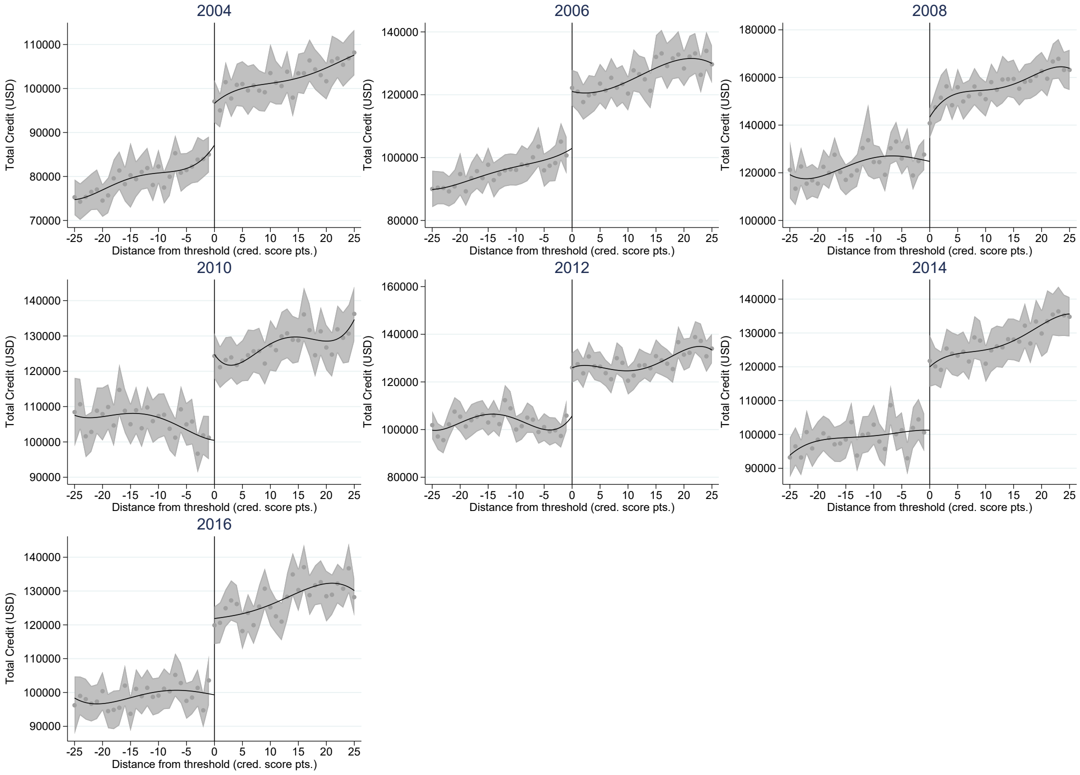


Figure C3: Credit score density smoothness around the thresholds over time

C.2 Accounting for the Gerrymandering of Congressional Districts

Table C9: Shifts in vote margins for Democrat and Republican candidates in US House general elections with 15 credit score point bandwidth.

	Republican votes		Democrat Votes	
VARIABLES	(1)	(2)	(3)	(4)
Share Close Thresh		0.541***		-0.497***
		(0.163)		(0.170)
Share Close Thresh (Above)	0.357**		-0.301	
	(0.179)		(0.187)	
Share Close Thresh (Below)	0.766***		-0.738***	
	(0.218)		(0.225)	
Share China Import	0.006	0.006	-0.016	-0.016
	(0.016)	(0.016)	(0.016)	(0.016)
Share White	0.807***	0.806***	-0.787***	-0.786***
	(0.027)	(0.027)	(0.028)	(0.029)
Share Female (voting age)	-0.588**	-0.587**	0.699***	0.698***
	(0.259)	(0.258)	(0.259)	(0.259)
Observations	4,180,761	4,180,761	4,180,761	4,180,761
R-squared	0.503	0.503	0.478	0.477

Clustered SE at county x congressional district

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

Table C9: The sample is restricted to election years 2012, 2014, and 2016 to exclude districts subject to potentially strategic redistricting after the 2010 census.

Table C10: Shifts in Poole-Rosenthal DW-NOMINATE scores for Democrat, Republican, and all candidates in US House general elections with 15 credit score point bandwidth.

	Republican votes		Democrat Votes		All Candidates	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Share Close Thresh		-0.167		0.462*		0.319
		(0.104)		(0.269)		(0.390)
Share Close Thresh (Above)	-0.230*		0.446		0.253	
	(0.127)		(0.326)		(0.431)	
Share Close Thresh (Below)	-0.085		0.479		0.399	
	(0.147)		(0.335)		(0.491)	
Share China Import	0.028**	0.028**	-0.039	-0.039	0.098**	0.098**
	(0.012)	(0.012)	(0.032)	(0.032)	(0.038)	(0.038)
Share White	0.034	0.034	0.262***	0.262***	1.200***	1.200***
	(0.028)	(0.028)	(0.037)	(0.037)	(0.066)	(0.066)
Share Female (voting age)	0.291	0.292	-0.390	-0.390	-0.132	-0.132
	(0.234)	(0.234)	(0.352)	(0.352)	(0.590)	(0.590)
Observations	2,307,950	2,307,950	1,886,554	1,886,554	4,194,504	4,194,504
R-squared	0.020	0.020	0.134	0.134	0.329	0.329

Clustered SE at county x congressional district
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C10: The sample is restricted to election years 2012, 2014, and 2016 to exclude districts subject to potentially strategic redistricting after the 2010 census.