

Quality and selection in regulated professions*

Gaetano Basso

Banca d'Italia

Eleonora Brandimarti

University of Geneva

Michele Pellizzari

University of Geneva, CEPR and IZA

Giovanni Pica

Università della Svizzera Italiana, CSEF and LdA

March 16, 2021

Abstract

Entry in many occupations is regulated with the objective to screen out the least able producers and guarantee high quality of output. Unfortunately, the available empirical evidence suggests that in most cases these objectives are not achieved. In this paper we investigate entry into the legal profession and we document that such a failure is due to the combination of the incomplete anonymity of the entry exam and the intergenerational transmission of business opportunities. We use microdata covering the universe of law school graduates in Italy from 2007 to 2013 matched with their careers and earnings up to 5 years after graduation. Variation generated by the random assignment of the entry exam grading commissions allows us to identify the role of family ties in the selection process. We find that connected candidates, i.e. those with relatives already active in the profession, are more likely to pass the exam and eventually earn more, especially those who performed poorly in law school. When we simulate the process of occupational choice assuming family connections did not matter, we find that strong positive selection on ability would emerge.

Keywords: Occupational regulation, licensing, intergenerational mobility.

JEL Codes: J24, J44, J62

*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 to 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. Any errors or omissions are the sole responsibility of the authors.

Corresponding author: Michele Pellizzari, Institute of Economics and Econometrics, Geneva School of Economics and Management, University of Geneva, 40 Bd du Pont d'Arve, CH-1211, Geneva 4, Switzerland. Email: michele.pellizzari@unige.ch.

1 Introduction

Entry in many occupations is regulated with the objective to protect consumers by selecting only the most able producers into the market (Bryson and Kleiner, 2010; Friedman and Kuznets, 1945; Kleiner, 2000; Kleiner and Krueger, 2013). However, the available empirical evidence suggests that in most cases occupational regulation fails to achieve such a goal (Anderson, Brown, Kerwin, and Rees, 2020; Bryson and Kleiner, 2019; Kleiner, 2017). The robustness of this finding across different professions and institutional contexts is indeed surprising, considering that 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).¹ Nonetheless, the literature has devoted very little attention to understanding the reasons of the surprising and generalised failure of occupational regulations.²

In this paper we investigate entry into the legal profession in Italy and we uncover 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 seem to increase significantly with their GPA. For non-connected candidates, i.e. those without family ties among licensed lawyers, GPA in law school matters a lot for the probability of passing the licensing exam and only those with the highest grades fully catch up with the pass rates of the connected ones. In addition, we find that, all else equal, connected lawyers earn more than non-connected lawyers, especially at the lower end of the distribution of law school grades. Our analysis shows that such earnings advantage is substantially larger when young connected lawyers work in the same law firms of their family ties.

Due to the combination of these effects, positive selection on ability into the profession is very limited, at least when it is measured along the distribution of GPA. Among law school graduates in the lowest GPA decile, about 46% eventually become licensed lawyers compared to 50% in the highest decile. With the help of a simple model, we simulate the selection process under the assumption that family connections were unimportant, both in the probability of passing the exam and in the earnings process. Our results show that, in this hypothetical scenario, occupational licensing would indeed produce strong positive selection on ability: the incidence 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.

This is a very important result from the point of view of policy design, as occupational

¹Anderson et al. (2020) is the one single paper finding positive effects of introducing regulation on quality.

²Some even argue that, in a world of online transactions, the prevalence of consumer ratings might make licensing redundant (Farronato, Fradkin, Larsen, and Brynjolfsson, 2020).

licensing affects about 20% of workers in the European Union and up to 30% in the United States (Kleiner and Krueger, 2013; Koumenta and Pagliero, 2018).³

Interestingly, we also document that, conditional on parental education, family connections have no predictive power on any measure of human capital that is available in our data, most notably high school and law school grades. Such evidence suggests that, at least for Italian lawyers, human capital transmission within the family does not seem to be occupation specific.⁴ Our analysis shows that the interplay of poorly designed regulation and the intergenerational transmission of occupations can severely undermine the potential of licensing to generate the positive selection on ability it is designed to create.

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 are clearly not general enough to 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 a 18-months apprenticeship period and an entry exam, consisting 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 practice making it very difficult to attract new clients as a young lawyer, are all factors that may naturally lead to favouring young entrants in the market who already have some connections with established professionals.⁵

As further discussed in Section 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 prices and commercial practices are extremely common across many professions, especially liberal professions, and countries (Paterson, Fink, and Ogus, 2003; Pellizzari and Pica, 2010; Pellizzari, Basso, Catania, Labartino, Malacrino, and Monti, 2011; UK Office of Fair Trade, 2001). 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 non-connected ones, often regardless of quality (Broscheid and Teske, 2003; Cannings, Montmarquette, and Mahseredjian, 1996; Chetty, Friedman, Saez, Turner, and Yagan, 2020).⁶

³The figure for the US combines proper licensing and certification.

⁴Lentz and Laband (1989) report similar findings for doctors in the US.

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

⁶Even though most US jurisdictions do not require formal apprenticeship periods in order to sit the bar exam,

In Italy, the bar exam is administered by 26 local districts, one per each court of appeal. Despite being common to many other countries (e.g. Germany, Canada, US), this decentralisation presents some peculiarities that offer the possibility to exploit useful variation for identification purposes. Specifically, each year the written exams of each district are marked by the commission of a different randomly selected district. Given substantial variation in grading standards across, this setup allows us to identify the process of selection via the exam separately from other mechanisms, such as the endogenous choice of starting the apprenticeship period and expectations of future earnings.

Our data combine university administrative records covering the universe of all law school graduates between 2007-2013 with the lists of all officially 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 derive information on family background, apprenticeship and earnings.

A crucial element of our analysis is the measurement of ability and connectedness. Our preferred proxy of professional ability is the GPA in law school and we also have information on high school grades, that we use as a proxy of general ability. Regarding connections, we build on a now rather extended literature using surnames and we code graduates as connected if their surnames appear at least once in the local register and among lawyers who obtained their license at least 25 years before the year of their (presumed) first attempt at the bar exam (Angelucci, De Giorgi, Rangel, and Rasul, 2010; Brollo, Kaufmann, and La Ferrara, 2017; Buonanno and Vanin, 2017; Güell, Mora, and Telmer, 2015; Güell, Pellizzari, Pica, and Mora, 2018).⁷ Of course, we transparently acknowledge that these are only imperfect proxies and we include extensive robustness checks to investigate the implications of measurement error for our main results (see Sections 7.1 and 7.2 for details).

We are not the first to look at the relationship between occupational regulation and quality of producers and/or output. Contrary to the predictions of the basic theory, most papers in this branch of the literature find no or even negative effects: our main contribution is to uncover potential mechanisms that could explain this surprising and rather robust finding.

Already Carroll and Gaston (1981) in their exploratory analysis 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,

the character and fitness requirement for admission is 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. Spain is considering shifting its admission process to one very similar to the Italian system.

⁷For robustness, we also experiment with the total number of times one’s surname appears in the local register.

Marier, Park, and Wing (2016) could not find any detectable improvement in the quality of health services when licensing regulations for nurses were 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, Ju, Kapor, and Yu (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 introduction of licensing requirements for midwives over time and across US states and find significant reductions in maternal and infant mortality.⁹

A recent advancement in this literature is Kleiner and Soltas (2019), 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 find 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.

Compared to some of these studies, our analysis is limited to the quality of service providers, or input quality. Measuring output quality can be an insurmountable task for many occupations and, although theoretically possible, it is hard to imagine situations in which input and output quality would move in opposite directions.

This paper is tightly connected to the literature on the intergenerational transmission of occupations. Already Lentz and Laband (1989) and Laband and Lentz (1992) 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, Roine, and Waldenström (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 very same employers. More directly related to our work, a recent literature documents sizeable intergenerational correlations

⁸Wanchek (2010) and Wing, Langelier, Continelli, and Battrell (2005) also investigate occupational regulations for dentists but do not focus on quality.

⁹Deyo, Hoarty, Norris, and Timmons (2020) also look at quality, but rather indirectly by studying the implications on crime and health of licensing massage therapists.

of professional affiliations in Italy (Aina and Nicoletti, 2018; Bamieh and Cintolesi, 2020; Mocetti, 2016; Mocetti and Roma, 2020; Mocetti, Roma, and Rubolino, 2018; Raitano and Vona, 2018). Compared to these papers, we link the intergenerational transmission of occupations to the effectiveness of licensing regulations by directly addressing selection and quality of professionals.

Many other papers have looked at occupational licensing in a variety of professions, but without a specific focus on quality. Most of these studies document an increase in costs for consumers and profits or rents for incumbent professionals. This is the case for driving school in France (Avrillier, Hivert, and Kramarz, 2010), lawyers (Pagliero, 2010, 2011), barbers (Thornton and Weintraub, 1979; Timmons and Thornton, 2010), and radiologists (Timmons and Thornton, 2008) in the USA.

The rest of the paper is organized as follows. Section 2 describes the institutional setup of the legal profession in Italy. Section 3 presents our main data sources and how we combine them. The model that guides our empirical investigation is introduced in Section 4. The empirical implementation of the model and the results are discussed in Section 5. In Section 6 we present counterfactual simulations allowing us to quantify the role of various mechanisms in the process of selection into the legal profession. Section 7 contains a large battery of robustness checks. Section 8 concludes.

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 then passing an entry exam.

The exam is organised by the courts of appeal, the second layer of the Italian judiciary system.¹⁰ There are 26 such courts in the country, approximately one 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. The first day they have to write an opinion on a civil case, the second day on a penal case and the third day they have to prepare a judiciary act. 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.

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

The written tests are graded anonymously by the commission of a randomly chosen district. The randomisation is performed by the Ministry of Justice and is clustered within 5 groups of districts of similar sizes. The randomisation is also designed to avoid pairing, namely two commissions grading each other. As an example, in 2019 the group of the biggest 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

For a variety of reasons, ranging from differences in social capital to differences in local labour market conditions, commissions in different parts of the country apply different standards to the correction of the exam papers. 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.

Figure 1 shows the average pass rates at the entry exam (both written and oral) over the period 2004-2012 over all districts. On average 34% of candidates eventually pass, but there are very large differences across districts, ranging from 27% in Turin (TO) to 50% in Palermo (PA).

Every year several thousand candidates attempt the written exam: on average during the period of our analysis, over 30,000 candidates participated every year. Hence, the grading process takes a long time, usually around 6 months, with some variation both over time and across districts. The written exam takes place on the same dates for districts, normally in the first half of December. In most cases, results are published sometime during the summer, and successful candidates are then admitted to the oral exam. The interviews happen in alphabetical order, starting from a randomly drawn letter. Each district draws its own letter and starts the oral examinations as soon as the results of their written exams are available, independently of the other districts. Each interview takes usually around one hour and candidates are notified the outcome at the end. In most districts the calendar of interviews spans all the months between September and December and, eventually, the entire process is completed only a few days before the new round of written exams begins.¹²

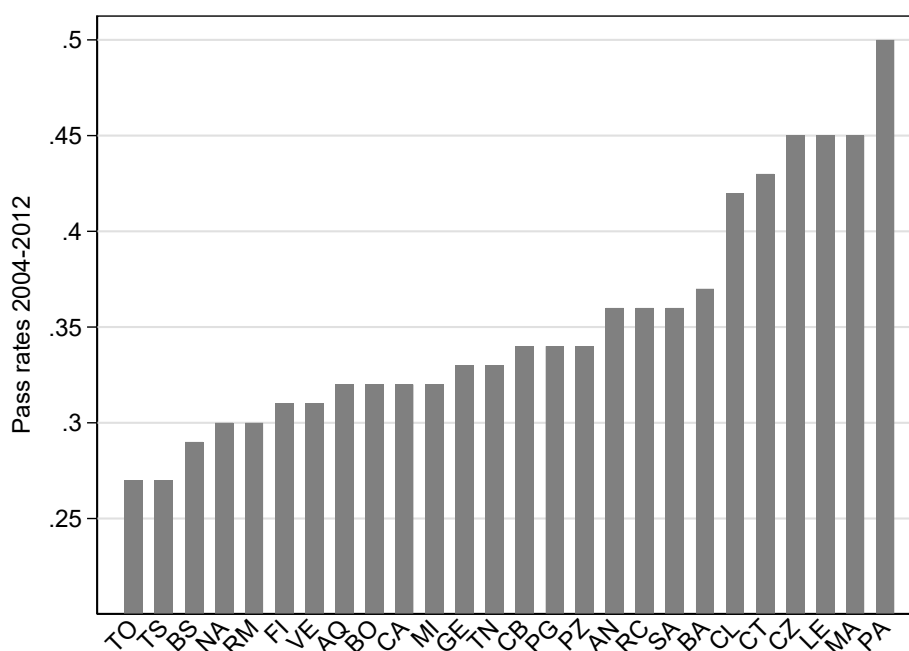
Given the 18 months of practice and the length of the examination process, young lawyers obtain their license approximately 2.5-3 years after graduation, unless they fail the exam (either the written or the oral part), in which case the process takes substantially longer. The entry exam can be retaken any number of times.¹³ Figure 2 summarizes the entry process into the profession, from the moment of graduation to the final occupational outcome. Candidates who

¹¹For 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.

¹²It is common for candidates who successfully passed the written exam, but are waiting to take the oral part to enroll in the written exam of the following year and be ready to sit it in case they would fail the interview.

¹³However, the completion of the 18-month training period is valid only for 5 years.

Figure 1: Pass rates at the bar exam



Source: Buonanno and Pagliero (2018).

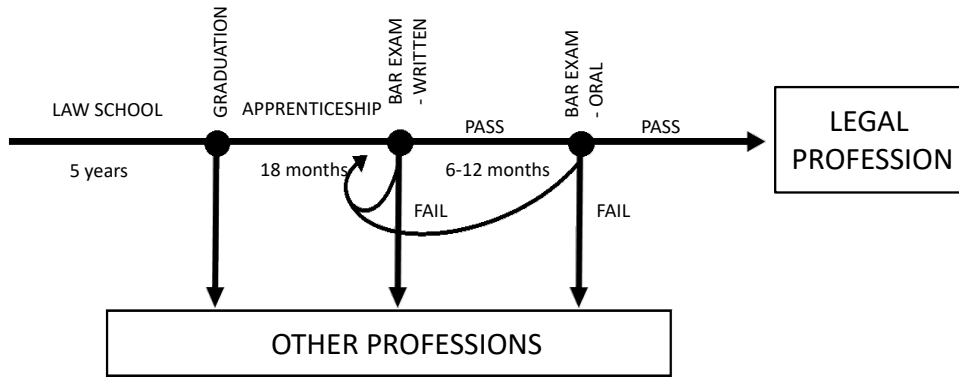
fail the exam can either retake or, in many cases, choose to enter a different occupation, often as legal consultants in private firms.

Candidates who successfully pass both the written and the oral exam 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. 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.¹⁴ 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 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. They nominate the local exam commissions which are composed of five members: three lawyers, nominated by the local bars, one (retired) judge and one university professor. The president of the commission must be chosen among the three lawyers. The local commissions

¹⁴The 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 2: Timeline of the licensing process for Italian lawyers



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 commission with the same composition is created at the Ministry of Justice, with the responsibility to prepare the questions of the written tests, defining the general criteria for grading and evaluating the oral exams and overseeing the entire examination process. All commissions are redone every year.

2.1 International comparison

The overall structure of the licensing process 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.¹⁵

The characteristics of the law degrees which give access to the vocational training are usually also highly uniformed across countries. Within the European Union, agreements exist allowing 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 study at the tertiary level.

¹⁵One 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.

The subsequent period of professional apprenticeship is usually organized in collaboration between universities and the state, with slight differences between 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.¹⁶

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 is commonly produced by attending a post-graduate law school (including both academic and vocational training), with entrance through a competitive examination. Germany requires two state-administered exams before becoming a lawyer: a first one after university which allows successful candidates to qualify for two years of compulsory training period (*Referendariat*), and a second one after its successful completion. 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).¹⁷ All these countries then require passing a state-administered exam, which upon successful completion allows an individual to practice law.¹⁸

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

Given the information we collected, we observe that most countries broadly follow a three-

¹⁶Most 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, just as in Italy.

¹⁷Japan’s vocational training takes place after sitting a nationally administered exam, which has the lowest success rate in the world, around 22%.

¹⁸The 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.

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, most notably the role of the inter-generational transmission of occupation, is likely to be present in many other contexts.

3 Data and descriptive evidence

We combine various data sources allowing us to follow several cohorts of Italian law school graduates over the first 5 years after graduation, observing, in particular, their occupational destinations and wages at the end of the process.

The starting point is an administrative dataset covering (almost) the entire universe of university graduates in Italy.¹⁹ This 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 after graduation and collects information about the students' backgrounds, opinions about the university experience and expectations about their professional careers. Then, students are interviewed again at one, three and five years since graduation to collect information about their labour market status. Almost all students fill in the survey at graduation, which is administered as part of the process to obtain their diplomas. The response rates of the other surveys are very high: 80% at one year, 75% at three years and 70% five years after graduation.

For this study we focus on students who graduated from a law school between 2007 and 2013. Before 2007, the follow-up surveys were only administered to those who graduated in the summer session (that is, about one third of graduates) and data on only 49 out of the 76 participating universities was available, whereas the 5-year post graduation surveys for students graduated after 2013 have not yet been released.

From the administrative records we observe high school type and marks, the university the students graduated from, GPA, graduation grades, age and gender. Important survey information includes 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. It is the responsibility of the local associations to publish and maintain the lists of licensed professionals in their

¹⁹All 50 public universities offering law degrees are included. Five of the nine private universities offering law degrees are included.

jurisdictions and most of the associations make them available on their websites. We have collected all of them during the period from November 2017 to January 2018. The registers contain information on the names, surnames and unique fiscal codes of all the associates. We use this information to match them with our main dataset of law school graduates, allowing us to identify those who eventually entered the legal profession and where.²⁰

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.²¹ Given the very high cost of entering the profession, we expect this to happen very 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 at 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 whether and how frequently their surnames appear in the local register (or in others). For lawyers, we define the local district as the one in which we observe them for the first time, whereas for non-lawyers we use the register corresponding to the location of the university from which they graduate. Of course, this is an imperfect measure of connectedness. There can be family 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 holds surnames that are very infrequent. Hence, the probability that any two individuals with the same surname are linked by some family tie is extremely high.²² It remains possible, however, that we fail to capture some connections because they do not share the same surname. In Section 7.2, we simulate various scenarios of mismeasurement to show that, under most assumptions, the degree of error must be very large to overturn our main findings, at least from the qualitative viewpoint.

In addition, from the public registers we also observe the professional coordinates of each lawyer - postal addresses, emails and phone/fax numbers - allowing us to identify those who are likely working in the same law firms. More specifically, we assume that any two lawyers

²⁰The matching has been performed for us by AlmaLaurea and we only have access to the matched anonymised version of the final dataset.

²¹It 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.

²²By using some very restrictive assumptions (i.e. that the size of the average household is equal to 3), (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.

reporting the same phone number or fax address or postal address in the register work in the same firm.²³

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.²⁴

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 7.4 we replicate all our estimates to show that the results are largely unaffected by this sample selection.

Table 1 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. Legal studies attract over 60% of female students and a little majority of them eventually end up not practicing as a licensed professional. The simple descriptive statistics suggest some minor positive selection on ability into the profession, both looking at high school and university grades.²⁵ 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 entering the legal profession: 88% of them do 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.

We conclude this section presenting some descriptive evidence on the selection of lawyers

²³Postal addresses, emails and phone/fax numbers are treated in strictly anonymous format. AlmaLaurea simply recorded all these variables with unique and anonymous alphanumeric codes. Ideally, one would like to match co-workers on the basis of phone and fax numbers. This is necessarily more precise and less likely to be prone to errors than using postal addresses because multiple firms might be located in the same building. We manually checked (by searching their websites on the internet) the largest resulting studios from matching on postal addresses to confirm that indeed the coordinates refer to a single firm. However, many younger lawyers do not provide fax numbers and 9.97% of those who provide a phone number give a cell phone number. In the end, we are able to retrieve matchable 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%), so we opt for matching on addresses. However, information on phone, fax and email is used to validate this matching, and we conclude that misclassification is a very minor problem.

²⁴We 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 quite reasonable approximation, especially regarding the adult population.

²⁵Notice that, 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 attributed via a common national exam.

Table 1: Descriptive statistics

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

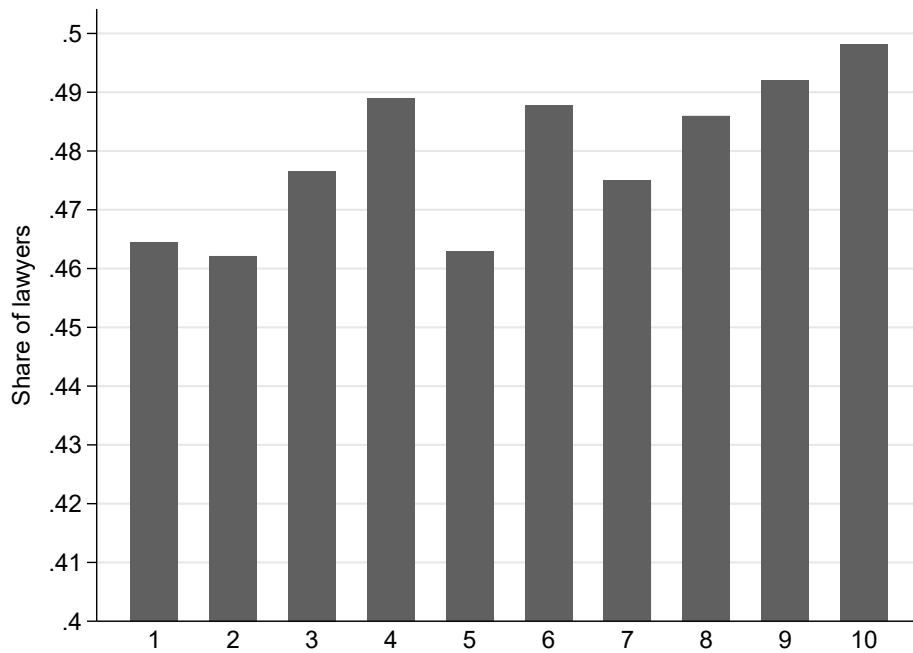
^f At least one parent with a university degree.

^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).

Figure 3: Share of lawyers by ability

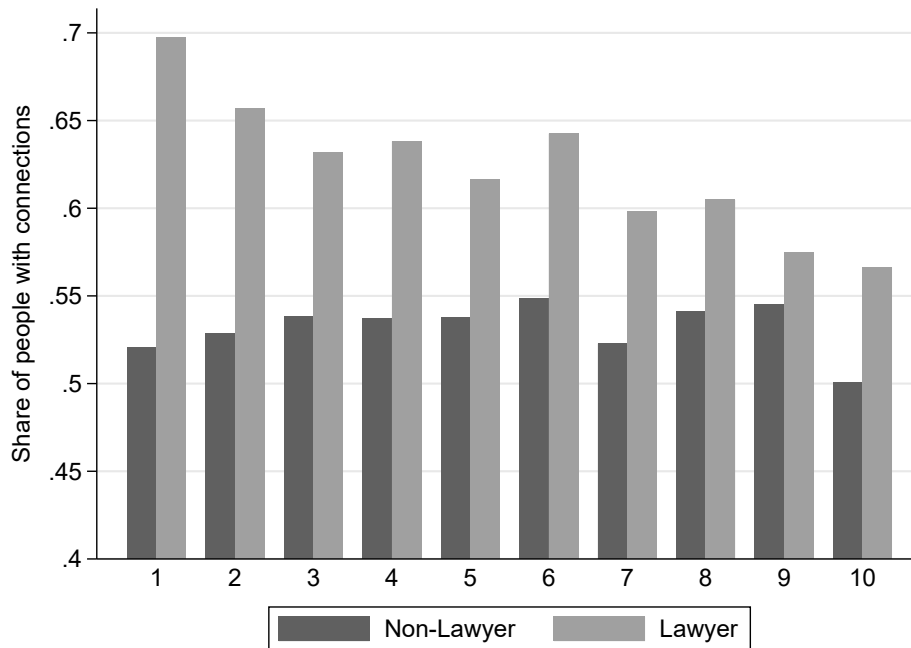


into the profession and the role of family connections. Figure 3 plots the share of law school graduates in our sample by decile of the distribution of GPA. This plot shows that some positive selection on ability does take place, but it is quite limited. About 46% of graduates in the first and second deciles access the profession and this share barely reaches 50% at the very top of the distribution.

Figure 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. Among 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 the non lawyers.

Taken together, Figure 3 and Figure 4 suggest that family connections might interfere with the selection process and explain, at least in part, the mild positive selection on ability that we detect in the data.

Figure 4: Share of connected individuals by occupation and ability



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 2.1).

The model consists of a sequential processes, mimicking the scheme in Figure 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 attempt the entry exam work in some other non-regulated occupation whereas those who successfully pass it become lawyers and generate earnings from professional practice. The following paragraphs describe how we model each of the steps, starting from earnings and moving backward in the sequence of events.

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:

$$Y_i = \begin{cases} y^L(A_i, S_i, N_i, G_i, X_i) + u_i^L & \text{if working as a lawyer} \\ y^0(A_i, S_i, N_i, G_i, X_i) + u_i^0 & \text{if working as a non-lawyer} \end{cases} \quad (1)$$

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. We measure 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.

4.2 Compulsory apprenticeship and bar exam

After graduation, individuals make occupational choices. Those who choose the legal profession have to 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 describing how we model the probability of passing the bar exam. We assume that the overall performance at the exam is a function 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.):

$$p(A_i, S_i, G_i, X_i) + \epsilon_i \quad (2)$$

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 2 we have documented the large heterogeneity in grading standards across districts (see Figure 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 usual 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 bar 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 not fully exogenous 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:

$$p(A_i, S_i, G_i, X_i) + \epsilon_i \geq t(R_r, N_i) \quad (3)$$

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 the family's socioeconomic status.²⁶ The intuition is that affluent parents are better able to support their children during this relatively long period with no or very 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.²⁷ 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) 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:

$$P[\epsilon_i \geq t(R_r, N_i) - p(A_i, S_i, X_i)] [y^L(A_i, S_i, N_i, G_i, X_i) - y^0(A_i, S_i, N_i, G_i, X_i)] + v_i \geq C(W_i) \quad (4)$$

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

²⁶This is actually very close to reality, as only in very few cases apprentices receive some salary.

²⁷This follows a relatively standard definition of social groups that is also adopted by the Italian National Statistical Institute (ISTAT) (ISTAT, 2017).

Eventually, the probability of becoming a lawyer can be computed as the product between the probability of starting an apprenticeship (equation (4)) and the probability of passing the bar exam, conditional on having started an apprenticeship (equation (3)).

4.3 Human capital formation

One important innovation of our data is the availability of a relatively good (although certainly not perfect) proxy of the agents' ability in legal matters, namely GPA in law school. We model the accumulation of this dimension of occupation specific human capital as follows:

$$S_i = s(A_i, N_i, G_i, W_i, X_i) + e_i \quad (5)$$

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 accumulation of occupation specific skills. In addition, we also allow connectedness to affect the accumulation of specific human capital, as one can acquire occupational skills from parents or other relatives with experience of the profession.

We assume here that all skills are fixed over our observation period. More specifically, we assume that the accumulation of generic skills is completed by the end of high school and the accumulation of specific skills is fixed at graduation from law school. These assumptions reflect the information available in the data, where we only observe proxies of individual ability at the end of high school and at graduation. Nevertheless, in Section 7.1, 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 (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.²⁸ At a minimum, equation (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 the implications of relaxing it in Section 7.1.

²⁸Perhaps the one variable that might be the most problematic is W_i . 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.

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: taken together and interpreted through the lenses of the model, they also allow running 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 6). The main results are produced using the restricted sample of observations that are available for all equations (see Table 1), thus avoiding compositional issues when comparing results across equations. In the robustness checks of Section 7.4 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) 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.

5.1 Human capital formation

We start with equation (5), which describes the process of occupation specific human capital formation. We assume linearity and we estimate it by simple OLS:

$$S_i = \beta_0 + \beta_1 A_i + \beta_2 N_i + \beta_3 G_i + \beta_5 W_i + \beta_4 X_i + e_i^S \quad (6)$$

Results are reported in Table 2. Perhaps not surprisingly, law school GPA is positively associated with both high school graduation marks and parental education. More importantly for the purpose of our paper is the lack of a statistically significant association between law school GPA and our indicator of connectedness with the legal profession. Our data does not support the notion that occupation specific human capital is transmitted within the family. If anything the results in Table 2 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.

Table 2: Occupation-specific human capital

Dep. variable= GPA ^a	(1)	(2)	(3)	(4)
High school grade ^b	0.402*** (0.006)	0.402*** (0.006)	0.402*** (0.006)	0.402*** (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*** (0.011)	0.098*** (0.011)	0.098*** (0.011)	0.098*** (0.011)
1=graduate parent ^e	0.126*** (0.011)	0.126*** (0.011)	0.125*** (0.011)	0.125*** (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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Given the importance of this finding, the following columns of Table 2 investigate whether it might be due to non-linearities in the relationship between GPA and connectedness. Column 2 categorises connections in three broad groups: no connections (the baseline), few connections (1 to 3) and many connections (4 or more). Column 3 and 4 further look at the linear number of connections and its square. None of the specifications seems to even remotely point 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 to perspective clients. 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 an study exchange abroad. In Table 3 we report the estimates of regression equations like equation (6) but with each of these indicators as dependent variables.

Table 3: Additional measures of human capital

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

^a Dummy equal to 1 if the respondent holds an internationally-recognized language certificate 5e.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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

While students with higher high school grades are more likely to hold certifications of proficiency in both foreign languages and computer skills, and are also more likely to volunteer, the coefficient on connectedness shows up as positive and significant at 10% only for foreign languages. The point estimates of the coefficients on connectedness for computer skills, volunteering and study exchanges are all non significant and very small.

Overall, the evidence in this section suggests that, at least for lawyers in Italy, the accumulation of occupation specific human capital within the family is very limited or even absent.

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 (3), which describes the probability of passing the exam (both written and oral), and equation (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.

The estimation samples are, however, different. While for the likelihood of starting an apprenticeship we can use all graduates, when we look at the probability of passing the bar exam we need to restrict the sample to those who actually took 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 specification for the probability of doing an apprenticeship and the probability of passing the bar exam:

$$P(T_i = 1|Z_i) = \Phi\{\theta_0^T + \theta_1^T S_i + \theta_2^T N_i + \theta_3^T (S_i \times N_i) + \theta_4^T A_i + \theta_5^T G_i + \theta_6^T W_i + \theta_7^T X_i\} \quad (7)$$

$$P(L_i = 1|T_i = 1, Z_i) = \Phi\{\theta_0^L + \theta_1^L S_i + \theta_2^L N_i + \theta_3^L (S_i \times N_i) + \theta_4^L A_i + \theta_5^L G_i + \theta_7^L X_i\} \quad (8)$$

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 $\{A_i, S_i, N_i, G_i, W_i, X_i\}$. 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 4.

We find that GPA matters for both passing the exam and deciding to undertake the apprenticeship period, while connections only for the former. The interaction of these two terms is

Table 4: Probabilities of apprenticeship and exam

	Probability of	
	doing an apprenticeship $P(T_i = 1 Z_{ir})$	passing the exam $P(E_i = 1 T_i = 1, Z_{ir})$
GPA ^a	0.026*** (0.004)	0.014** (0.006)
1=connections ^b	0.007 (0.005)	0.045*** (0.009)
GPA × [1=connections]	-0.008* (0.004)	-0.027*** (0.007)
High school grade ^c	-0.014*** (0.002)	-0.006 (0.004)
1=female	0.012*** (0.004)	-0.033*** (0.007)
1=graduate parent ^d	0.005 (0.005)	-0.006 (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

^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.

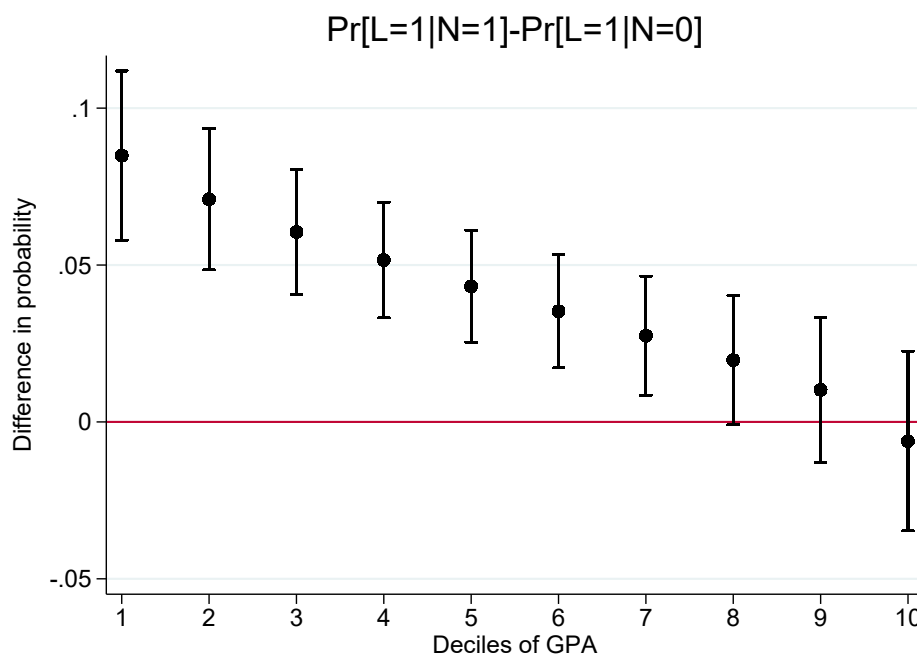
^e At least one parent employed as professional, entrepreneur or manager.

^f Fixed effects for the district of exam correction.

All regressions include fixed effects for university, district, and year of graduation. Probit coefficients are reported. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

negative and significant only in the probability of passing the bar exam. This suggest that GPA is more important for non-connected candidates than connected ones. To further investigate this important issue and get a sense of the magnitudes, Figure 5 shows the difference in predicted probabilities of passing the bar exam for connected and non-connected candidates by deciles of the distribution of GPA.

Figure 5: Difference in predicted pass rates by ability and connections



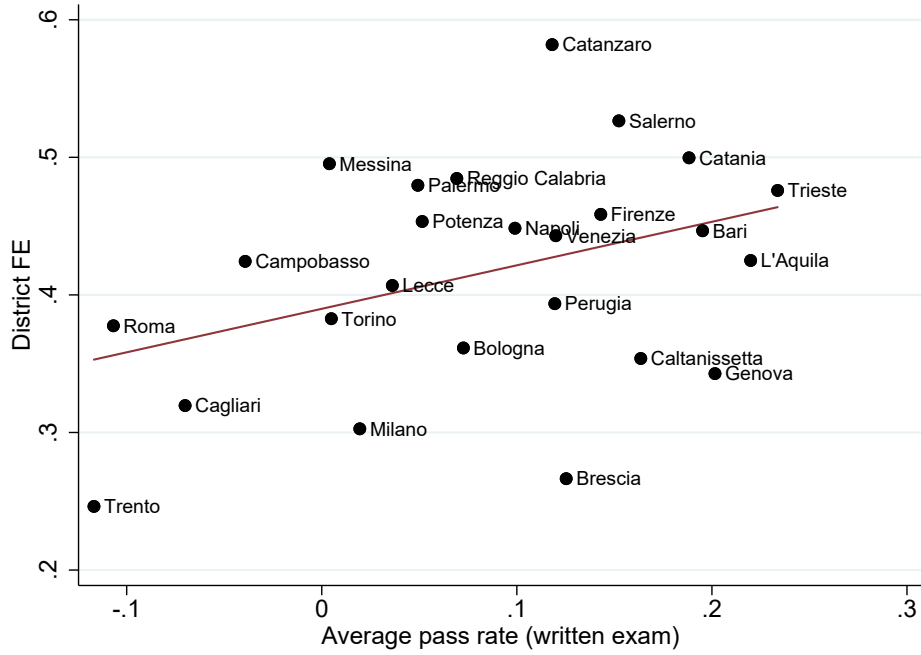
Note: Predictions based on estimates Table 4, column 2.

Connected candidates are systematically more likely to pass the exam, especially at low levels of GPA, where the difference is almost 10 percentage points. Interestingly, GPA matters a great deal for non-connected candidates and very little for connected ones. As GPA goes up, the gap between non-connected and connected candidates narrows down until it finally disappears from the eighth decile onward.

To the extent that GPA captures professional ability, these results suggest that family connections can heavily distort the selection process into the profession. Notice that, having controlled for ability and parental human capital, it is unlikely that the results in Table 2 and Figure 5 are generated by transmission of abilities within the family, which already appeared to be rather unimportant from Table 2.

In Figure 6 we show the fixed effects for the district of exam correction, as estimated from equation(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 to the average pass rates at the written

Figure 6: District fixed effects and pass rates at the written exam.



Note: Fixed effects for the district of exam correction from Table 4, column 2.

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. 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.

5.3 Earnings

As a last step, we estimate equation (1), once again assuming linear functional forms:

$$y_i^L = \alpha_0^L + \alpha_1^L S_i + \alpha_2^L N_i + \alpha_3^L (S_i \times N_i) + \alpha_4^L A_i + \alpha_5^L G_i + \alpha_6^L X_i + \nu_i^L \quad (9)$$

$$y_i^0 = \alpha_0^0 + \alpha_1^0 S_i + \alpha_2^0 N_i + \alpha_3^0 (S_i \times N_i) + \alpha_4^0 A_i + \alpha_5^0 G_i + \alpha_6^0 X_i + \nu_i^0 \quad (10)$$

where y_i^J is the log of the monthly earnings that individual i self-reported in the 5-year post-graduation survey. Importantly, we do not want to drop individuals who report zero earnings, mostly because lawyers are self-employed and, thus, sometimes may have zero earnings. More generally, 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 in the non-legal sector. In order to avoid dropping the observations with zero earnings when taking logs, we simply add one to all records.

Equations (9) and (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 and passing the bar exam:

$$P(L_i = 1|Z_{ir}) = \Phi [\theta_0 + \theta_1 S_i + \theta_2 N_i + \theta_3 (S_i \times N_i) + \theta_4 A_i + \theta_5 G_i + \theta_6 W_i + \theta_7 R_r + \theta_8 X_i] \quad (11)$$

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 and we are aware that its role might be more questionable than that of r .²⁹

Result are reported in Table 5. We find that 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 grade is 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.³⁰

Having relatives in the profession is also associated with higher earnings, but only for those working as licensed lawyers, which is consistent with our interpretation of what this variable should capture. In addition, the interaction of GPA and connectedness is negative for lawyers (and non-significant for non-lawyers), suggesting that occupation specific ability might presumably be less important when one has facilitated access to a portfolio of potential clients via family ties.

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

²⁹We have also experimented a version of the model without W_i and the estimates are robust. Results are available upon request.

³⁰Consistent with a large body of empirical evidence, we also find very significant gender gaps in earnings.

Table 5: 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** (0.039)	0.133*** (0.046)	0.073*** (0.015)
1=connections ^b	0.173*** (0.057)	-0.096 (0.084)	0.122*** (0.024)
GPA × [1=connections]	-0.122** (0.045)	0.026 (0.056)	-0.081*** (0.018)
High school grade ^c	0.063*** (0.024)	0.088*** (0.031)	-0.032*** (0.010)
1=female	-0.627*** (0.044)	-0.658*** (0.057)	-0.067*** (0.018)
1=graduate parent ^d	0.045 (0.043)	0.031 (0.054)	-0.030 (0.020)
<i>Exclusion restrictions:</i>			
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

^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.

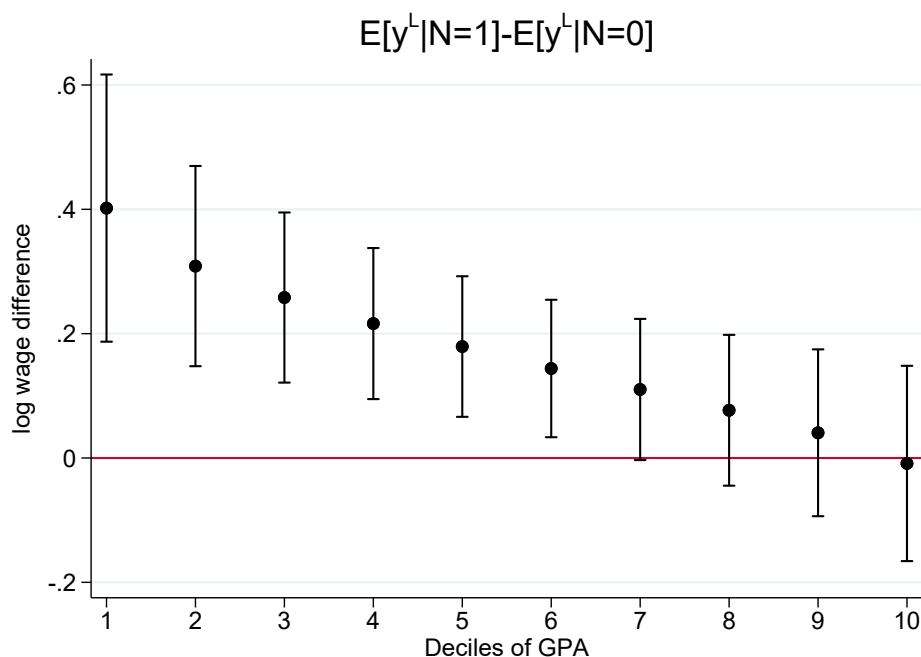
All regressions include fixed effects for university, district and year of graduation. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

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 5. 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 5, Figure 7 shows the predicted difference in log earnings in the legal profession between connected and non-connected lawyers along the distribution of GPA. In the bottom decile of the distribution, connected lawyers earn 40% more than their non-connected colleagues and it is only towards the very top of the distribution that this difference becomes statistically insignificant.

Figure 7: Predicted (log) wage differences between connected and non-connected lawyers by ability



Notes: Predictions based on the estimates of Table 5, columns 1 and 2. The vertical bars represent 95% confidence intervals.

One possible interpretation of this effect is that non-connected lawyers, especially at the beginning of their careers, have a very hard time accessing clients. Having relatives who can partially share their portfolio of clients and perhaps expand it, might represent a very significant advantage. This interpretation is also consistent with the very strict regulations concerning professional practice. In Italy, like in many other countries in continental Europe, professional

associations impose codes of conduct that regulate, among other things, commercial practice. For example, it is often prohibited to approach clients who are already served by another professional and, until recently, commercial advertisement was considered to be contrary to the “dignity of the profession”. In addition, the code of conduct indicates price floors.³¹ Being unable to lower prices and to advertise their services, it is extremely difficult for new entrants in the market for legal services to attract clients.

5.3.1 Working with relatives

In this Section, we present additional evidence suggesting that the effect of connections on earnings in the legal profession is substantially larger when young lawyers work in the exact same law firm as their connections. We do not formally incorporate the process of selection into law firms into our model of Section 4 because its identification would require an additional exclusion restriction, which we do not have. Hence, we present these results merely as descriptive evidence in support of our interpretation of the main findings.

Using our proxy of law firms based on professional coordinates (see Section 3 for details), we find that about 7% of the lawyers in our data work with relatives. We augment equation (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 indicator. We then estimate the following equation:

$$y_i^L = \alpha_0^L + \alpha_1^L S_i + \alpha_2^L N_i + \alpha_3^L (S_i \times N_i) + \alpha_4^L A_i + \alpha_5^L G_i + \alpha_6^L X_i + \alpha_7^L F_i + \alpha_8^L (F_i \times S_i) + \nu_i^L \quad (12)$$

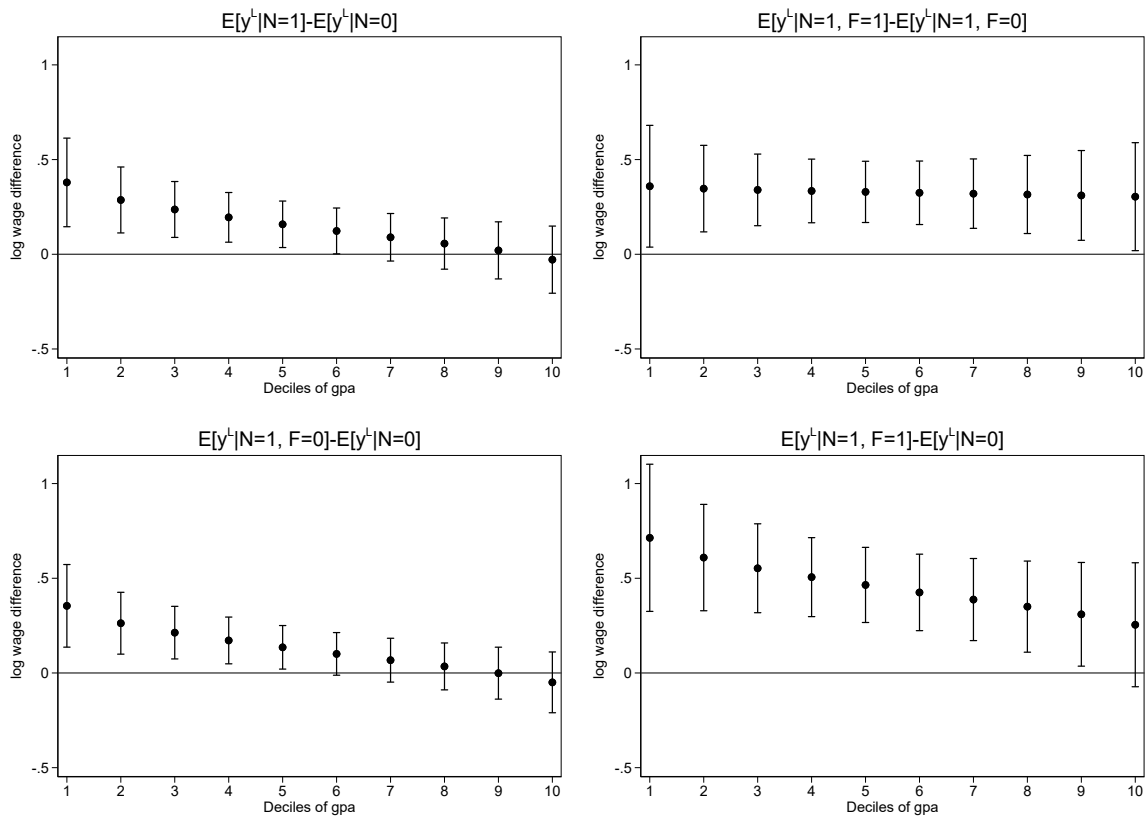
For brevity, we only report results graphically. Figure 8 follows the same logic as Figure 7, but it extends the comparison to lawyers working in the same firms as their connections or in others. For completeness, Panel A replicates under the specification of equation (12), the same analysis of Figure 7, namely the comparison of the average earnings of connected and non-connected 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 8 focuses exclusively on connected individuals and compares the average earnings of those working with relatives with the others. Throughout the distribution of GPA, young lawyers working in the same law firm as some relative earn around 4% more than colleagues who, despite having family connections with the profession, do not work with them in the same firm.

Panel C compares connected lawyers who do not work with their relatives against non-connected colleagues. The resulting figure is very similar to Panel A, with a difference of 2.5-3

³¹Some of these regulations were reformed recently in Italy but, in the daily practice of the profession, they remain strongly present.

Figure 8: Predicted (log) wage differences by occupation, ability, connectedness and firm type.



Notes: In Panel A, we show the baseline result (wage differences between connected vs. non-connected 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 non-connected lawyers. Finally, Panel D reports the wage differences between connected lawyers working with relatives and non-connected ones.

The vertical bars represent 95% confidence intervals.

percentage points in the lowest deciles that vanishes approximately above the median.

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

Overall, the results in this section show that when connected lawyers work in the same firm as their family tie their earning advantage over non connected colleagues increases substantially and remains significant across almost the entire distribution of GPA.

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 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 5 suggests that the effect of connections on the probability of passing the exam is stronger at lower levels of GPA. To the extent that academic performance in law school captures occupation specific ability, this could generate negative selection or, at least, mitigate positive selection. 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.

Theoretically, there also exists a third channel, namely the differential accumulation of human capital, especially occupation specific human capital. However, in Section 5.1 we show that, at least in the case of Italian lawyers, the data does not seem to indicate that connected and non-connected individuals might be accumulating human capital differently. Hence, we will disregard this channel in our simulation exercise.³²

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 (3) and (4), respectively, and in Section 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 (5) we estimated it as described in equation (7), which does not allow disentangling the role of connectedness on earnings and the probability of passing the bar exam.

³²Of course, our framework can be applied to other settings and, if the human capital channel appeared to be important, the simulation exercise could easily be extended to incorporate it.

Hence, we go back to the definition of the probability of doing an apprenticeship presented in the theoretical Section (4), equation (4). First, we use the estimates of equations (3), (9) and (10) to compute the expected earnings premium in the legal profession (conditional on doing the apprenticeship):

$$\widehat{E}[\Delta y_i | Z_{ir}] = \widehat{y}^L(A_i, S_i, N_i, G_i, X_i) - \widehat{y}^0(A_i, S_i, N_i, G_i, X_i) \quad (13)$$

Then, we re-estimate the probability of the apprenticeship directly from its theoretical definition in equation (4):

$$\begin{aligned} \widehat{P}(T_i = 1 | Z_{ir}) &= P \left[v_i < \widehat{P}(L_i = 1 | T_i = 1, Z_{ir}) \widehat{E}[\Delta y_i | Z_{ir}] - \widehat{\theta}_6^T W_i \right] = \\ &\Phi \left[\widehat{P}(L_i = 1 | T_i = 1, Z_{ir}) \widehat{E}[\Delta y_i | Z_{ir}] - \widehat{\theta}_6^T W_i \right] \end{aligned} \quad (14)$$

where $\Phi(\cdot)$ is the cumulative density of the standard normal distribution. Notice that we had already assumed normality of v_i in Section 5, so there are no additional assumption in equation (14).³³

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

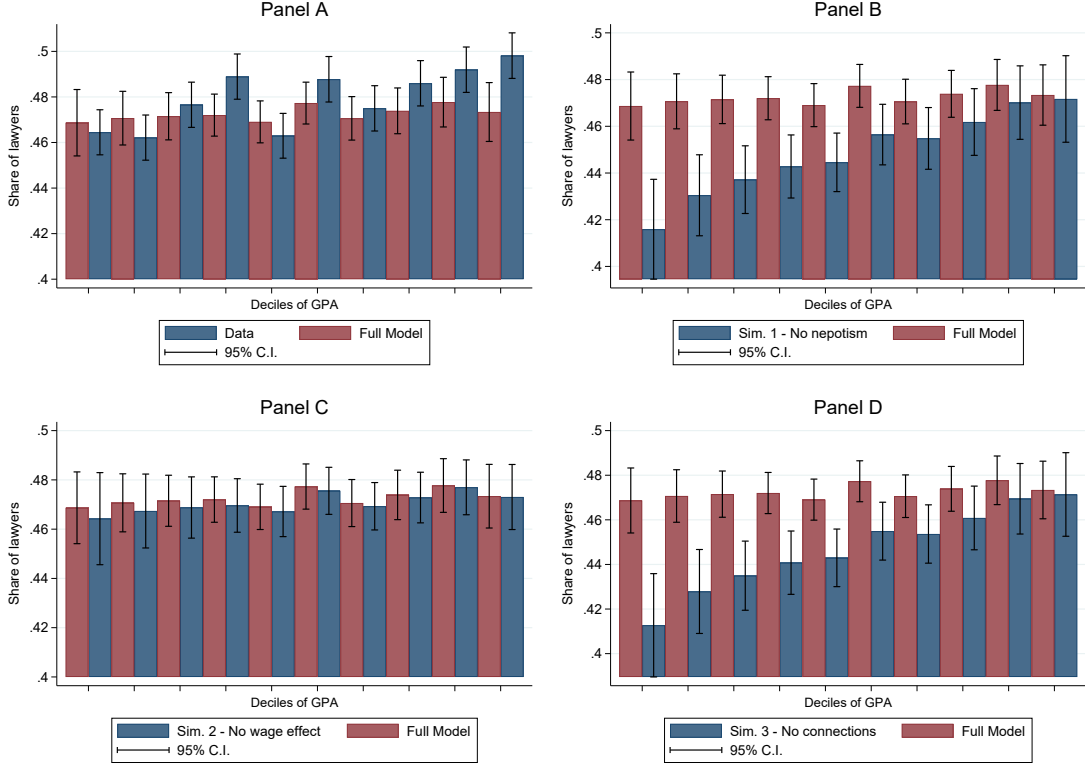
$$\begin{aligned} \widehat{P}(L_i = 1 | Z_{ir}) &= \widehat{P}(L_i = 1 | T_i = 1, Z_{ir}) \widehat{P}(T_i = 1 | Z_{ir}) \\ &= \widehat{P}(L_i = 1 | T_i = 1, Z_{ir}) \\ &\quad \Phi \left[\widehat{P}(L_i = 1 | T_i = 1, Z_{ir}) \widehat{E}[\Delta y_i | Z_{ir}] - \widehat{\theta}_6^T W_i \right] \end{aligned} \quad (15)$$

Panel A of Figure 9 compares the average estimated probability of being a lawyer from equation (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 higher 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 on ability that is observed in the data.

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.

³³To 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.

Figure 9: Counterfactual simulation exercises



Notes: The Figure reports the results of the counterfactual simulation exercises described in Section 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.

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

$$\hat{P}_B(L_i = 1|Z'_{ir}) = \hat{P}(L_i = 1|T_i = 1, Z'_{ir}, N_i = 0) \Phi \left[\hat{P}(L_i = 1|T_i = 1, Z'_{ir}, N_i = 0) \hat{E}[\Delta y_i|Z_{ir}] - \hat{\theta}_6^T W_i \right] \quad (16)$$

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:

$$\begin{aligned} \widehat{P}_C(L_i = 1|Z'_{ir}) &= \widehat{P}(L_i = 1|T_i = 1, Z_{ir}) \\ &\Phi \left[\widehat{P}(L_i = 1|T_i = 1, Z_{ir}) \widehat{E}[\Delta y_i|Z'_{ir}, N_i = 0] - \widehat{\theta}_6^T W_i \right] \end{aligned} \quad (17)$$

Contrary to the previous analysis, now the effect is much more limited and the predicted shares of lawyers by deciles are very similar under this scenario and the original model.

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

$$\begin{aligned} \widehat{P}_D(L_i = 1|Z'_{ir}) &= \widehat{P}(L_i = 1|T_i = 1, Z'_{ir}, N_i = 0) \\ &\Phi \left[\widehat{P}(L_i = 1|T_i = 1, Z'_{ir}, N_i = 0) \widehat{E}[\Delta y_i|Z'_{ir}, N_i = 0] - \widehat{\theta}_6^T W_i \right] \end{aligned} \quad (18)$$

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 low-ability lawyers, while high ability 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. Overall, without connections there would be fewer lawyers on the market with comparatively higher ability, and the channel through which connections impact the probability of becoming a lawyer is through the probability of passing the exam, rather than through potentially higher future earnings and a greater network base.

We believe that the existing structure of the bar exam in Italy does allow for nepotistic practices to emerge. Presumably this happens due to two factors: (i) the important role of incumbent lawyers in the process and (ii) the very 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.

7 Robustness checks

In this Section, we present several robustness checks to complement our main analysis. Section 7.1 and Section 7.2 investigate the implications of measurement error in our measures of

professional ability and family connections, respectively. In Section 7.3, 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 7.4 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 5.

7.1 Measurement error in ability

In our main analysis, we measure professional ability with GPA in law school and we have already acknowledged several times that this is only an imperfect proxy of true ability.

To understand the implications of this mismeasurement for our main results, it is useful to consider the following simplified version of our main equations:

$$Y_i = \psi_0 + \psi_S S_i^* + \psi_N N_i + \omega_i \quad (19)$$

where Y_i is a generic outcome (earnings in equations (9)-(10), passing the bar exam or participation in apprenticeship in equations (7)-(8)), S_i^* is the true unobservable professional ability and N_i is an indicator of family connections, as in our usual notation. For the sake of simplicity and conciseness, we abstract from all complications associated with functional forms and multiple equations (most of our main equations are not linear and we often estimate multiple equations simultaneously). We also abstract from the interaction of S_i^* and N_i that is instead present in all equations of our main exercise and from measurement error in N_i , which we discuss in the next section.³⁴

We are interested in the consistent estimation of parameters ψ_S and ψ_N when S^* is not observable and we replace it with an observable proxy S , law school GPA in our setting. Following standard results, the proxy variable S is normally required to satisfy the following two assumptions.

1. The proxy variable S needs to be uncorrelated with the error term of the original equation, ω . This assumption implies that if we were to estimate equation (19) augmented with S_i , the coefficient on the proxy variable would converge to zero. In other words, GPA should carry no additional information if true ability could be observed.
2. True ability S^* needs to be uncorrelated with connections N_i , once conditioning on the proxy variable S . Formally, if we could run the regression:

$$S_i^* = \tau_0 + \tau_N N_i + \tau_S S_i + \mu_i \quad (20)$$

³⁴It is easy to show that the main results we derive in this section remain valid when we consider also the interaction of S_i^* and N_i and we allow both to be mismeasured.

the estimator of τ_N should converge to zero.

It is easy to show that, under these two assumptions, simply replacing S^* with S in equation (19) allows identifying ψ_N consistently, whereas the coefficient on the proxy variable converges to $\psi_S\tau_S$.

Let us now discuss the validity of the two assumptions in our specific setting and speculate on the sign and magnitude of τ_S , which is the key parameter to understand the implications of our proxy strategy for the estimation of the effects of professional ability on our outcomes of interest.

The first assumption is perhaps the easiest to defend. It is difficult to imagine reasons why GPA in law school should matter for outcomes such as passing the bar exam or earnings, above and beyond the effect of true ability. Even signalling effects of university grades are very unlikely to materialise, especially when considering the bar exam, whose written test is marked anonymously and where the commissioners of the oral part are not given information on the candidates' performance at university. Given our rich set of controls, including university and district fixed effects, it is unlikely that such signalling effects persist in earnings five years after graduation.

The second assumption is more compelling. In our specific setting it would be invalidated if professional human capital would be acquired also outside formal education. Beyond law school, we can think of two settings in which lawyers could accumulate professional skills: the family and the professional apprenticeship. In Section 5.1, we have already provided evidence that learning of legal practice within the family seems to be limited.

Regarding the apprenticeship, it is indeed very reasonable to imagine that young lawyers learn professional skills during this period but, unfortunately, indicators of ability measured after apprenticeship do not exist, as only a certificate of completion – a compulsory requirement for sitting the bar exam – is issued. Hence, we have investigated this issue in an extension of our theoretical framework in which agents exert effort to accumulate occupation-specific human capital in two steps: first, in law school and then during the apprenticeship period (see Appendix A for a full sketch of the model). In this extended model we assume, rather naturally, that the production function of human capital displays complementarity between the human capital accumulated in law school and the effort exerted during the apprenticeship period. We show that, under this assumption, measures of professional ability at the end of law school are strongly positively correlated with similar measures at the end of apprenticeship. In the terminology of equation (20), this implies that τ_S is positive. At a minimum, proxying ability with GPA preserves the ranking of lawyers based on true ability.

This result would be reversed 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 apprenticeship.³⁵ 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.

To understand the plausibility of this scenario, we investigate 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 lawyers and apprentices of the Veneto region, for whom we observe both income and the law firm in which they are employed, on top of their surnames and demographic information.³⁶ We use the licensed lawyers and their earnings in this dataset to estimate firm fixed effects – which we interpret as a measure of the quality of law firms – and we match them to the apprentices. Hence, we can check whether connected apprentices sort into firms of higher quality compared to the non-connected apprentices.³⁷

Figure 10 shows the resulting distribution of the firm fixed effects, separately for connected and non connected apprentices. As the distributions largely overlap, the exercise supports the notion that connected and non connected apprentices spend the apprenticeship period in law firms of similar quality. That is, there is no evidence that connected individuals receive additional training from high quality lawyers during their apprenticeship.

Overall, the theoretical and empirical results presented in this section suggest that using GPA in law school as a proxy of professional ability is unlikely to bias the estimated effects of connections. In addition, we show that our measurement strategy preserves both the sign of the effects of true ability and the ranking of agents.

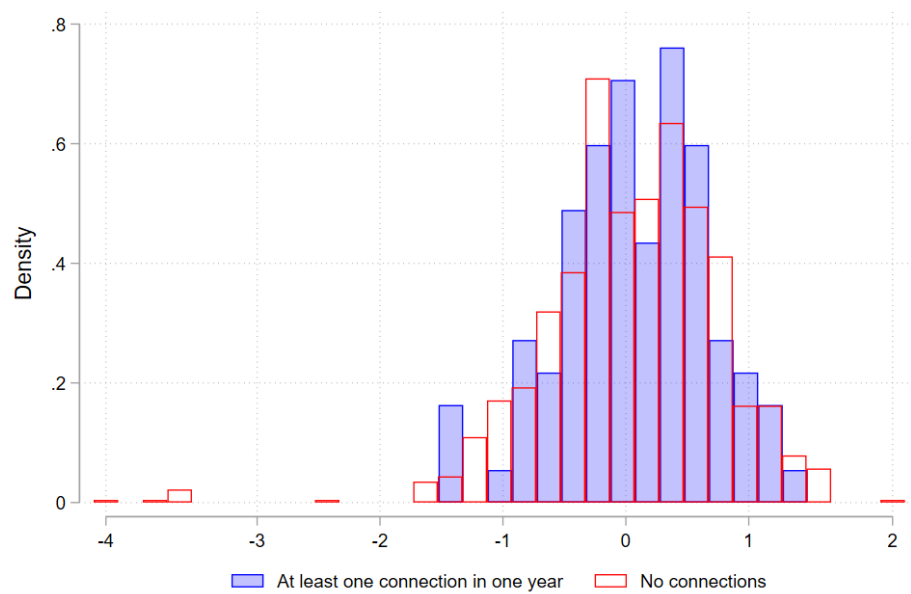
It remains to discuss whether τ_S is larger or smaller than 1, i.e. whether we over- or underestimate the effects of ability. Under the additional reasonable assumption that GPA is a noisy but unbiased measure of ability (i.e., the noise is not systematically related to any unobservable skills or other individual characteristics of the individual), it is easy to show that $\tau_S < 1$. Notice that the both results in Figure 10 and those reported in Section 5.1 suggest the absence of significant correlation between GPA and connections. Thus, our data appears to be consistent with the assumption that GPA is a noisy and unbiased measure of ability. Eventually, our results are best interpreted as lower bounds of the true effects of ability on the various outcomes that

³⁵Notice that a small degree of substitutability would not suffice to undo the positive relationship between law school performance and final professional ability.

³⁶Veneto registers, excluding that of Venice, were collected in the context of a previous project focusing on this region (see Pellizzari and Pica (2010)).

³⁷To estimate firms fixed effects, we restrict the analysis to licensed lawyers (i.e., we exclude apprentices) and regress (log) income on 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, year dummies and firms fixed effects. Unfortunately, for this 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.

Figure 10: Distribution of the fixed effects of the firms where apprentices spend the apprenticeship period, separately for connected and non connected apprentices



Note: Firms 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).

we analyse.

7.2 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, Hendren, Kline, and Saez, 2014; Raitano and Vona, 2018). 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-children links. Using surnames is subject to the mis-classification errors discussed above, but it allows capturing some 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 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.³⁸

Table 6 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 (8) (compare with column 2 of Table 4). In column 2, we replicate the same estimation, but we randomly recode 1% of the connected individuals as non-connected and we randomly take an equal number of non-connected individuals recoding them as connected.³⁹ The following

³⁸Results for the other equations confirm the findings in this section and can be obtained upon request.

³⁹We also experimented with other forms of recoding, such as recoding a given share of the connected and of the non-connected, 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

Table 6: Measurement error in connections and the probability of passing the bar exam

	Percentage of randomly re-assigned connections ^a					
	0% [1]	1% [2]	5% [3]	10% [4]	20% [5]	30% [6]
GPA ^b	0.042** (0.016)	0.036** (0.016)	0.035** (0.016)	0.026* (0.015)	0.017 (0.015)	0.006 (0.015)
1=connections ^c	0.128*** (0.026)	0.113*** (0.025)	0.097*** (0.023)	0.060*** (0.022)	0.020 (0.020)	0.010 (0.019)
GPA × [1=connections]	-0.079*** (0.019)	-0.070*** (0.019)	-0.069** (0.019)	-0.055** (0.019)	-0.040** (0.018)	-0.022 (0.018)
High school grade ^d	-0.017 (0.011)	-0.017 (0.011)	-0.017 (0.011)	-0.018 (0.011)	-0.018* (0.011)	-0.018* (0.011)
1=female	-0.091*** (0.019)	-0.091*** (0.019)	-0.092*** (0.019)	-0.092*** (0.019)	-0.094*** (0.019)	-0.094*** (0.019)
1=graduate parent ^e	-0.018 (0.019)	-0.017 (0.019)	-0.015 (0.019)	-0.013 (0.019)	-0.012 (0.019)	-0.012 (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 non-connected individuals is randomly assigned to being 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 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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

columns perform the same exercise with more and more random re-classifications. 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 and tend to disappear only when we reclassify large shares of individuals, i.e., more than 20%.

Table 7: 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*** (0.039)	0.230*** (0.039)	0.218*** (0.038)	0.219*** (0.038)	0.182*** (0.036)	0.172*** (0.035)
1=connections ^c	0.173*** (0.057)	0.174*** (0.056)	0.158*** (0.053)	0.156*** (0.049)	0.102** (0.045)	0.043 (0.042)
GPA × [1=connections]	-0.122*** (0.045)	-0.112** (0.045)	-0.094** (0.044)	-0.097** (0.043)	-0.041 (0.042)	-0.026 (0.041)
High school grade ^d	0.063*** (0.024)	0.063** (0.024)	0.061** (0.024)	0.061** (0.024)	0.061** (0.024)	0.061** (0.024)
1=female	-0.627*** (0.044)	-0.627*** (0.044)	-0.627*** (0.044)	-0.628*** (0.044)	-0.629*** (0.044)	-0.630*** (0.044)
1=graduate parent ^e	0.045 (0.043)	0.046 (0.043)	0.050 (0.043)	0.052 (0.043)	0.054 (0.043)	0.056 (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 non-connected individuals is randomly assigned to being 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 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 7 we reproduce the same exercise for the earnings equations (9) and (10), with equation (11) completing the switching regression model. For brevity, we only report results for earnings in the legal profession and, similarly to Table 6, 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 can affect the magnitude of our most important estimates, it seems unlikely that it is large enough to overturn their qualitative message.

in results exclusively to mis-measurement.

7.3 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 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*).⁴⁰ For non-lawyers we compute the growth rates of earnings from the official Italian Labour Force Survey, 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.⁴¹

Figure 11 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 12 reproduces the exercise by redefining expected (log) earnings as the full stream of discounted future earnings over 30 years of experience:

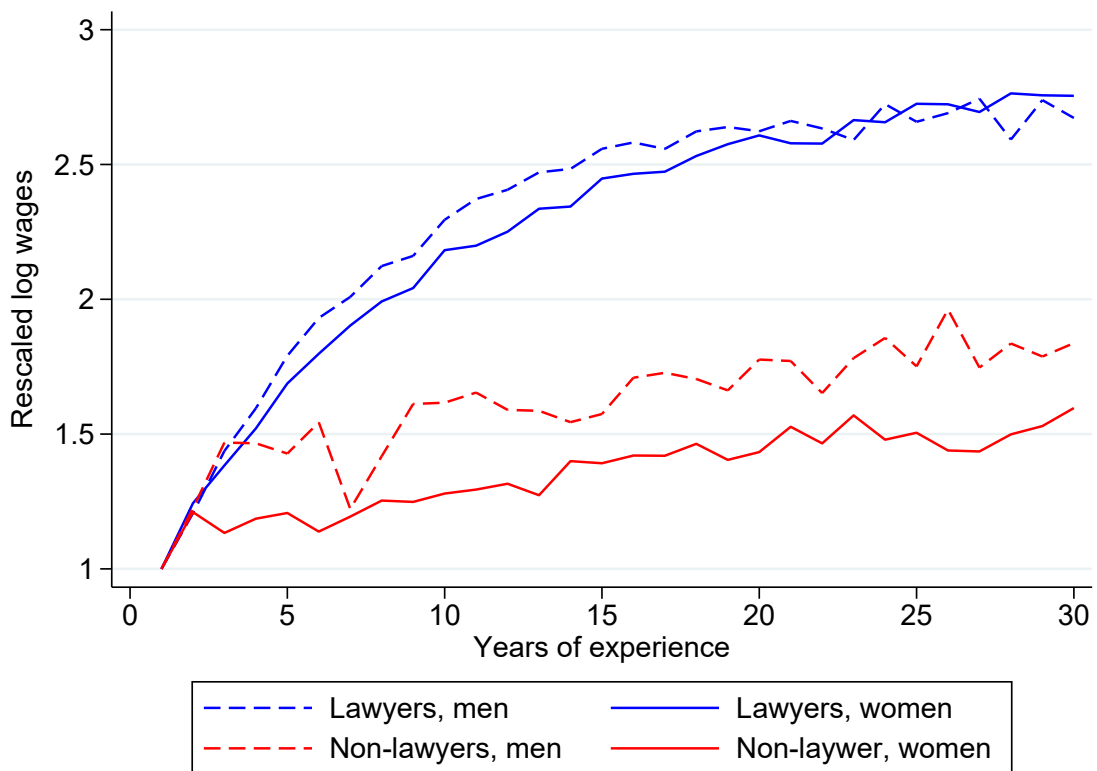
$$\widehat{E} [\Delta y_i | Z_{ir}] = \widehat{E} \left[(y_i^L + \sum_{e=1}^{30} \gamma_{ie}^L) - (y_i^0 + \sum_{e=1}^{30} \gamma_{ie}^0) | Z_{ir} \right] \quad (21)$$

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

⁴⁰We 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 21.

⁴¹Each yearly labour force survey is a representative cross-section of the Italian population. We estimate the experience profiles in the same way they are estimated for lawyers (see footnote 40), 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 21.

Figure 11: Wage-experience profiles for lawyers and non-lawyers by gender



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

$$\widehat{E}[\Delta y_i | Z'_{ir}, N_i = 0].$$

Results are very similar to those in our main simulations of Section 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 bit 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 of higher ability, who already start off with higher earnings. One potential limitation is that the experience profiles might be different for connected and non-connected individuals and we, unfortunately, have no information about this particular issue. 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 non-connected 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 21, 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.

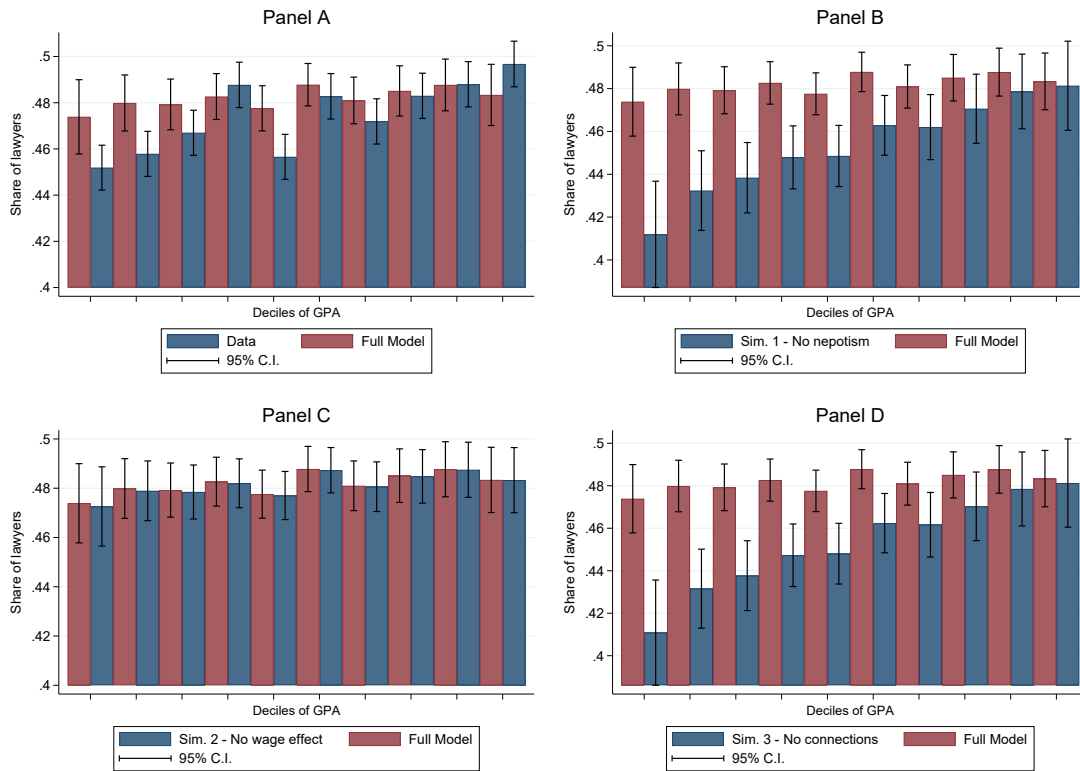
Eventually, and as in our baseline simulation, eliminating the role of connections both in earnings and in the exam yields very similar results as when eliminating them only in the exam (Panel D).

7.4 Sample composition

In our main analysis of Section 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 some doubts about the implications of sample selection. To clear such potential doubts, in this section we replicate our main estimates using the largest available sample for each equation separately.

Table 8 is the equivalent of Table 2 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, whereas most of the information needed to produce the estimates in Table 8 comes directly

Figure 12: 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 7.3. 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.

Table 8: Occupation specific human capital with largest possible sample

Dep. variable= GPA ^a	(1)	(2)	(3)	(4)
High school grade ^b	0.405*** (0.004)	0.405*** (0.004)	0.405*** (0.004)	0.405*** (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*** (0.000)	-0.002*** (0.001)
Number of connections ²	-	-	-	-0.000 (0.000)
1=female	0.106*** (0.008)	0.106*** (0.008)	0.106*** (0.008)	0.106*** (0.008)
1=graduate parent ^e	0.150*** (0.008)	0.150*** (0.008)	0.150*** (0.008)	0.150*** (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.

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

from administrative archives, where the issue of missing data is minimal.

Table 9 reproduces a similar exercise with reference to Table 4 in our main results of Section 5. The number of available observations is smaller than in the previous Table 8 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 non-connected individuals. Overall, our main results are confirmed.

Table 9: Probabilities of apprenticeship and exam with largest possible sample

	Probability of	
	doing an apprenticeship $P(T_i = 1 Z_{ir})$	passing the exam $P(E_i = 1 T_i = 1, Z_{ir})$
GPA ^a	0.128*** (0.016)	0.004 (0.013)
1=connections ^b	0.067** (0.027)	0.140*** (0.021)
GPA × [1=connections]	-0.034* (0.019)	-0.080*** (0.015)
High school grade ^c	-0.074*** (0.012)	-0.037*** (0.009)
1=female	0.053*** (0.021)	-0.084*** (0.016)
1=graduate parent ^d	0.035 (0.023)	-0.047*** (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, and year of graduation. Probit coefficients are reported. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Finally, in Table 10 we look at earnings, expanding the size of the sample as much as possible for each equation. Once again, results are extremely comparable to those reported in our main analysis.

Table 10: Lawyer and non-lawyer earnings with largest possible sample

	Lawyer earnings y_i^L	Non-lawyer earnings y_i^0	Selection $P(L_i = 1 Z_{ir})$
GPA ^a	0.250*** (0.040)	0.126*** (0.047)	0.054*** (0.013)
1=connections ^b	0.192*** (0.057)	-0.129 (0.089)	0.118*** (0.021)
GPA × [1=connections]	-0.130*** (0.046)	0.040 (0.058)	-0.087*** (0.016)
High school grade ^c	0.058** (0.024)	0.108*** (0.035)	-0.012 (0.009)
1=female	-0.629*** (0.045)	-0.658*** (0.056)	-0.094*** (0.016)
1=graduate parent ^d	0.049 (0.043)	0.026 (0.054)	-0.048*** (0.017)
<i>Exclusion restrictions:</i>			
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. Normalized 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 randomly assigned district of exam correction.

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

8 Conclusions

The available evidence indicates that occupational regulation almost invariably fails to improve the quality of professionals and the services they provide. Whereas this finding is quite 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.

Of course, our findings are specific to the context that we analyse, namely that of licensed lawyers in Italy, and they may not be easily generalized to all 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 the malfunction of the system is 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 have an interest in facilitating connected candidates and they might be able to do so lawfully and even unconsciously. 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 the complete abolition of the oral examination.

A number of important avenues remain open for future research. Among the most important ones we can mention the investigation of output quality. The ultimate aim of the regulation is guaranteeing the quality of the 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 or proxy input quality satisfactorily in this paper. Measuring output quality is even more challenging and Anderson et al. (2020) is the only paper we are aware of that does it. In addition, we acknowledge very transparently 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.

Bibliography

- Carmen Aina and Cheti Nicoletti. The intergenerational transmission of liberal professions. Labour Economics, 51:108 – 120, 2018.
- Mark Anderson, Ryan Brown, Charles Kerwin, and Daniel Rees. Occupational licensing and maternal health: Evidence from early midwifery laws. The Journal of Political Economy, Forthcoming, 2020.
- Manuela Angelucci, Giacomo De Giorgi, Marcos A. Rangel, and Imran Rasul. Family networks and school enrolment: Evidence from a randomized social experiment. Journal of Public Economics, 94(3-4):197–221, April 2010.
- Joshua D. Angrist and Jonathan Guryan. Does teacher testing raise teacher quality? Evidence from state certification requirements. Economics of Education Review, 27(5):483 – 503, 2008.
- Paul Avrillier, Laurent Hivert, and Francis Kramarz. 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, 2010.
- Omar Bamieh and Andrea Cintolesi. Intergenerational transmission in regulated professions and the role of familism. 2020.
- John Barrios. Occupational licensing and accountant quality: Evidence from the 150-hour rule. BFI Working Papers 2018-32, Becker Friedman Institute, 2018.
- Gaetano Basso. Barriers to enter the italian lawyers’ profession: An empirical surname-based approach. Master’s thesis, Bocconi University, 2009.
- Anders Bjorklund, Jesper Roine, and Daniel Waldenström. Intergenerational top income mobility in Sweden: Capitalist dynasties in the land of equal opportunity? Journal of Public Economics, 96(5):474–484, 2012.

- Fernanda Brollo, Katja Kaufmann, and Eliana La Ferrara. Learning about the enforcement of conditional welfare programs: Evidence from Brazil. CEPR Discussion Papers 11965, C.E.P.R. Discussion Papers, April 2017.
- Andreas Broscheid and Paul E. Teske. Public members on medical licensing boards and the choice of entry barriers. Public Choice, 114(3-4):445–59, 2003.
- Alex Bryson and Morris M. Kleiner. The regulation of occupations. British Journal of Industrial Relations, 48(4):670–675, December 2010.
- Alex Bryson and Morris M. Kleiner. Re-examining advances in occupational licensing research: Issues and policy implications. British Journal of Industrial Relations, 57(4):721–731, December 2019.
- Paolo Buonanno and Mario Pagliero. Occupational licensing, labor mobility, and the unfairness of entry standards. CEPR Discussion Papers 13076, Centre for Economic Policy Research, 2018.
- Paolo Buonanno and Paolo Vanin. Social closure, surnames and crime. Journal of Economic Behavior & Organization, 137(C):160–175, 2017.
- Kathy Cannings, Claude Montmarquette, and Sophie Mahseredjian. Entrance quotas and admission to medical schools: A sequential probit model. Economics of Education Review, 15(2):163–174, April 1996.
- Sidney L. Carroll and Robert J. Gaston. Occupational restrictions and the quality of service received: Some evidence. Southern Economic Journal, 47(4):959–976, 1981.
- Raj Chetty, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. Where is the land of opportunity? The geography of intergenerational mobility in the United States. The Quarterly Journal of Economics, 129(4):1553–1623, 2014.
- Raj Chetty, John N. Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. Income segregation and intergenerational mobility across colleges in the united states. Quarterly Journal of Economics, 2020.
- Miles Corak and Patrizio Piraino. The intergenerational transmission of employers. Journal of Labor Economics, 29(1):37–68, 2011.
- Darwyn Deyo, Blake Hoarty, Conor Norris, and Edward Timmons. Licensing massage therapists in the name of crime: the case of Harper v Lindsay. Journal of Entrepreneurship and Public Policy, Forthcoming, 2020.

- Thomas Dunn and Douglas Holtz-Eakin. Financial capital, human capital, and the transition to self-employment: Evidence from intergenerational links. Journal of Labor Economics, 18 (2):282–305, 2000.
- Chiara Farronato, Andrey Fradkin, Bradley Larsen, and Erik Brynjolfsson. Consumer protection in an online world: An analysis of occupational licensing. Working Paper 26601, National Bureau of Economic Research, January 2020.
- Milton Friedman and Simon Kuznets. Income from Independent Professional Practice. National Bureau of Economic Research, Inc, 1945.
- Maia Güell, José V. Rodríguez Mora, and Christopher I. Telmer. The Informational Content of Surnames, the Evolution of Intergenerational Mobility, and Assortative Mating. Review of Economic Studies, 82(2):693–735, 2015.
- Maia Güell, Michele Pellizzari, Giovanni Pica, and José V. Rodríguez Mora. Correlating social mobility and economic outcomes. Economic Journal, 128(612):F353–F403, 2018.
- Deborah Haas-Wilson. The effect of commercial practice restrictions: The case of optometry. The Journal of Law and Economics, 29(1):165–186, 1986.
- ISTAT. Definizione dei gruppi sociali e loro descrizione. Note to ch. 2, Rapporto Annuale ISTAT, 2017.
- Morris M. Kleiner. Occupational licensing. Journal of Economic Perspectives, 14(4):189–202, 2000.
- Morris M. Kleiner. The influence of occupational licensing and regulation. IZA World of Labor, pages 392–392, October 2017.
- Morris M. Kleiner and Alan B. Krueger. Analyzing the extent and influence of occupational licensing on the labor market. Journal of Labor Economics, 31(S1):173–202, 2013.
- Morris M. Kleiner and Robert T. Kudrle. Does regulation affect economic outcomes? The case of dentistry. The Journal of Law and Economics, 43(2):547–582, 2000.
- Morris M. Kleiner and Evan J. Soltas. A welfare analysis of occupational licensing in US states. NBER Working Papers 26283, National Bureau of Economic Research, 2019.
- Morris M. Kleiner, Allison Marier, Kyoung Won Park, and Coady Wing. Relaxing occupational licensing requirements: Analyzing wages and prices for a medical service. The Journal of Law and Economics, 59(2):261–291, 2016.

- Maria Koumenta and Mario Pagliero. Occupational regulation in the european union: Coverage and wage effects. British Journal of Industrial Relations, 57(4):818–849, 2018.
- Adriana D. Kugler and Robert M. Sauer. Doctors without borders? Relicensing requirements and negative selection in the market for physicians. Journal of Labor Economics, 23(3): 437–465, 2005.
- David Laband and Bernard F. Lentz. Self-recruitment in the legal profession. Journal of Labor Economics, 10(2):182–201, 1992.
- Bradley Larsen, Ziao Ju, Adam Kapor, and Chuan Yu. The effect of occupational licensing stringency on the teacher quality distribution. NBER Working Papers 28158, National Bureau of Economic Research, Inc, 2020.
- Hayne E. Leland. Quacks, lemons, and licensing: A theory of minimum quality standards. Journal of political economy, 87(6):1328–1346, 1979.
- Bernard F. Lentz and David Laband. Why so many children of doctors become doctors: Nepotism vs. human capital transfers. Journal of Human Resources, 24(3):396–413, 1989.
- Alex Maurizi. Occupational licensing and the public interest. The Journal of Political Economy, 82(2):399–413, 1974.
- Sauro Mocetti. Dynasties in professions and the role of rents and regulation: Evidence from italian pharmacies. Journal of Public Economics, 133:1–10, 2016.
- Sauro Mocetti and Giacomo Roma. Le professioni ordinistiche in italia: misure ed effetti della regolamentazione. 2020.
- Sauro Mocetti, Giacomo Roma, and Enrico Rubolino. Knocking on parents’ doors: Regulation and intergenerational mobility. Economic Working Papers 1182, Bank of Italy, 2018.
- Jacopo Orsini and Michele Pellizzari. Dinastie d’Italia: gli ordini tutelano davvero i consumatori? Università Bocconi Editore, 2012.
- Mario Pagliero. Licensing exam difficulty and entry salaries in the us market for lawyers. British Journal of Industrial Relations, 48(4):726–739, 2010.
- Mario Pagliero. What is the objective of professional licensing? evidence from the us market for lawyers. International Journal of Industrial organization, 29(4):473–483, 2011.
- Iain Paterson, Marcel Fink, and Anthony Ogus. Economic impact of regulation in the field of liberal professions in different Member States. Institute for Advanced Studies (IHS), 2003.

- Michele Pellizzari and Giovanni Pica. Liberalizing professional services: Evidence from Italian lawyers. Technical Report 372, Bocconi University, 2010.
- Michele Pellizzari, Gaetano Basso, Andrea Catania, Giovanna Labartino, Davide Malacrino, and Paola Monti. Family ties in licensed professions in Italy. Technical report, Fondazione Rodolfo Debenedetti, 2011.
- Michele Raitano and Francesco Vona. Nepotism vs specific skills : the effect of professional liberalization on returns to parental back grounds of Italian lawyers. Sciences po publications, Sciences Po, November 2018.
- Carl Shapiro. Investment, moral hazard, and occupational licensing. The Review of Economic Studies, 53(5):843–862, 1986.
- George J. Stigler. The Theory of Economic Regulation. Bell Journal of Economics, 2(1):3–21, Spring 1971.
- Robert J. Thornton and Andrew R. Weintraub. Licensing in the barbering profession. ILR Review, 32(2):242–249, 1979.
- Edward J. Timmons and Robert J. Thornton. The effects of licensing on the wages of radiologic technologists. Journal of Labor Research, 29(4):333–346, 2008.
- Edward J. Timmons and Robert J. Thornton. The licensing of barbers in the USA. British Journal of Industrial Relations, 48(4):740–757, 2010.
- UK Office of Fair Trade. Competition in Professions. OFT, 2001.
- Tanya Wanchek. Dental hygiene regulation and access to oral healthcare: Assessing the variation across the US states. British Journal of Industrial Relations, 48(4):706–725, 2010.
- Paul Wing, Margaret H. Langelier, Tracey A. Continelli, and Ann Battrell. 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), 2005.

A Appendix. Human capital accumulation and effort

To study how occupation-specific human capital accumulation during the compulsory apprenticeship impacts the likelihood of passing the bar exam we sketch a simple two-period model. We assume that individuals (with a certain predetermined level of general human capital A) accumulate occupation-specific human capital in two steps, first at law school and then during the apprenticeship. At each step, human capital accumulation requires effort and depends positively on the amount of human capital 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 favors 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:

$$S_2 = \Phi_2 A^{\gamma_1} (GPA)^{\gamma_2} e_2^{\gamma_3} \quad (22)$$

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 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:

$$\beta_2 \gamma_2 f'(S_2) \Phi_2 A^{(\gamma_1 + \beta_1 \gamma_2)} \Phi_1^{\gamma_2} \left[e_1^{\beta_2 \gamma_2 - 1} e_2^{\gamma_3} \right] = \delta_1 \quad (23)$$

$$\gamma_3 f'(S_2) \Phi_2 A^{(\gamma_1 + \beta_1 \gamma_2)} \Phi_1^{\gamma_2} \left[e_1^{\beta_2 \gamma_2} e_2^{\gamma_3 - 1} \right] = \delta_2 \quad (24)$$

Notice that the marginal productivity of effort in each period depends positively on the effort exerted in the other period. Dividing equation (23) by equation (24) we get:

$$\frac{e_2}{e_1} = \frac{\delta_1}{\delta_2} \frac{\gamma_3}{\beta_2 \gamma_2} \quad (25)$$

Notice the following:

1. Neither the Φ s, 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 (22) 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 (δ s) 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 (25) into equation (23) we get:

$$\beta_2 \gamma_2 f'(S_2) \Phi_2 A^{(\gamma_1 + \beta_1 \gamma_2)} \Phi_1^{\gamma_2} \left[e_1^{\beta_2 \gamma_2 - 1} e_2^{\gamma_3} \right] = \delta_1 \quad (26)$$

$$\beta_2 \gamma_2 f'(S_2) \Phi_2 A^{(\gamma_1 + \beta_1 \gamma_2)} \Phi_1^{\gamma_2} \left[e_1^{\beta_2 \gamma_2 - 1} \left(e_1 \frac{\delta_1}{\delta_2} \frac{\gamma_3}{\beta_2 \gamma_2} \right)^{\gamma_3} \right] = \delta_1 \quad (27)$$

$$\beta_2 \gamma_2 f'(S_2) \Phi_2 A^{(\gamma_1 + \beta_1 \gamma_2)} \Phi_1^{\gamma_2} \left[e_1^{\beta_2 \gamma_2 - 1 + \gamma_3} \left(\frac{\delta_1}{\delta_2} \frac{\gamma_3}{\beta_2 \gamma_2} \right)^{\gamma_3} \right] = \delta_1 \quad (28)$$

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):

$$e_1^* = \left[\frac{\beta_2 \gamma_2 \Phi_2 A^{(\gamma_1 + \beta_1 \gamma_2)} \Phi_1^{\gamma_2} \left(\frac{\delta_1 \gamma_3}{\delta_2 \beta_2 \gamma_2} \right)^{\gamma_3}}{\delta_1} \right]^{\frac{1}{1 - \beta_2 \gamma_2 - \gamma_3}} \quad (29)$$

$$e_2^* = e_1^* \frac{\delta_1 \gamma_3}{\delta_2 \beta_2 \gamma_2} \quad (30)$$

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 the two input be instead substitute, the result would be overturned. However, we view complementarity as a more plausible assumption. Thus, in our empirical analysis, the proxy for occupation-specific human capital, GPA at law school, likely incorporates the expected benefits of apprenticeships in high-quality law firms and, therefore, it does not provide a systematically biased proxy of occupation-specific human capital.