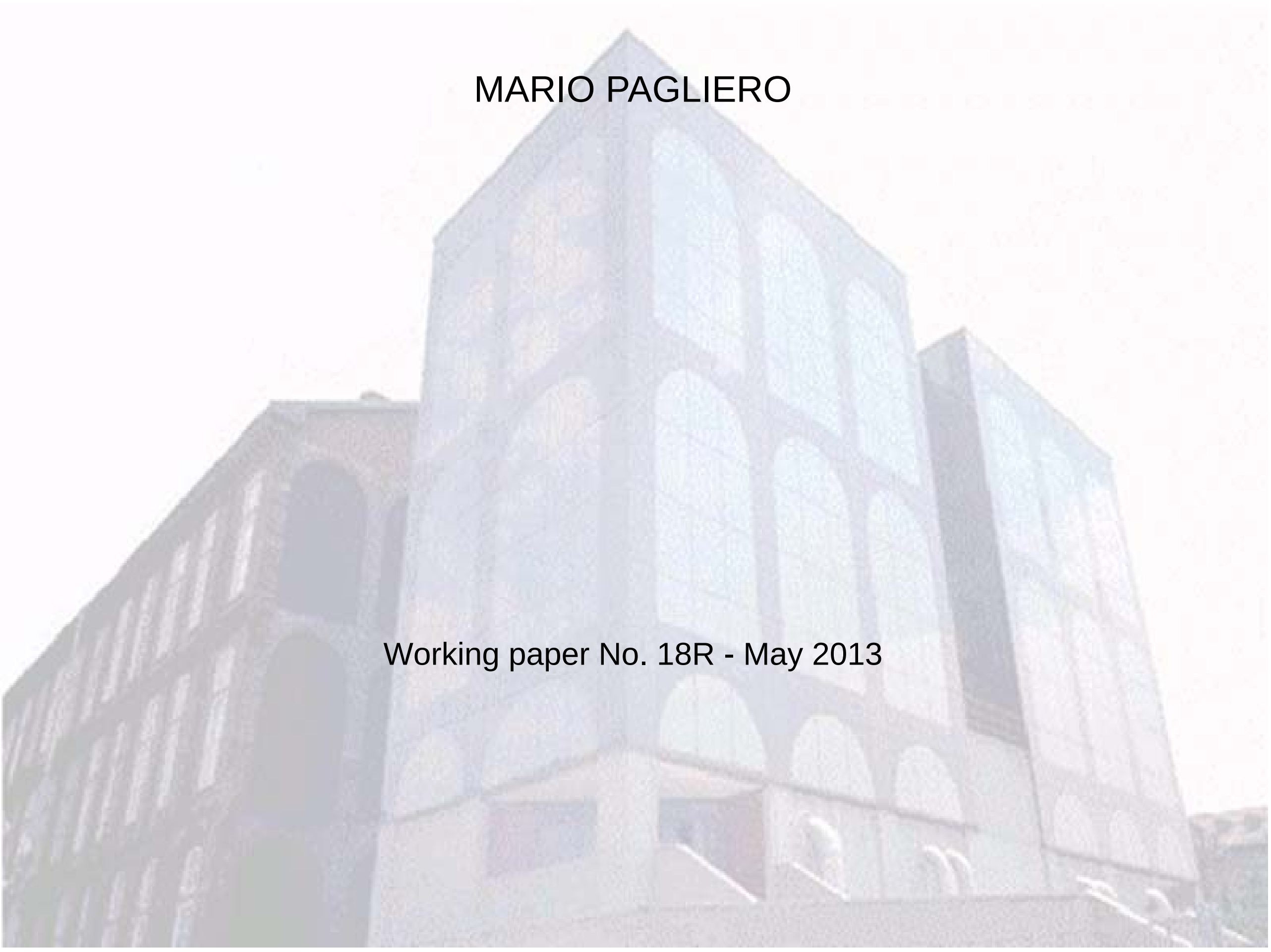
UNIVERSITÀ
DEGLI STUDI
DI TORINO

ALMA UNIVERSITAS
TAURINENSIS

Founded in 1404

DEPARTMENT OF ECONOMICS AND
PUBLIC FINANCE "G. PRATO"
WORKING PAPER SERIES

THE IMPACT OF POTENTIAL LABOR SUPPLY ON LICENSING EXAM DIFFICULTY



MARIO PAGLIERO

Working paper No. 18R - May 2013

The Impact of Potential Labor Supply on Licensing Exam Difficulty

Mario Pagliero*

February 2013

Abstract

Entry into licensed professions requires meeting competency requirements, typically assessed through licensing examinations. This paper explores whether the number of individuals attempting to enter a profession (potential supply) affects the difficulty of the entry examination. The empirical results suggest that a larger potential supply may lead to more difficult licensing exams and lower pass rates. This implies that licensing may partially shelter the market from supply shocks and limit the impact of policies targeted at increasing labor supply.

JEL: L4, L5, J44, K2.

Keywords: occupational licensing, minimum standards, entry regulation, legal market.

*University of Turin and Collegio Carlo Alberto; e-mail: mario.pagliero@unito.it. I would like to thank Cristian Bartolucci, Maristella Botticini, Francesca Cornelli, Pascal Courty, Christos Genakos, David Genesove, Mindy Marks, Aleksey Tetenov, seminar participants at the Collegio Carlo Alberto, European University Institute, London Business School, Universitat Pompeu Fabra, University of Amsterdam, University of Tilburg, and at the 2011 AEA meetings for helpful comments and discussions. Any remaining errors are those of the author alone.

1 Introduction

For an increasing number of occupations, people seeking to enter the profession must satisfy a number of requirements set by state licensing boards. This usually means passing a licensing examination and meeting educational, residency and moral character and fitness requirements. According to Kleiner (2000), over 800 occupations are licensed in at least one U.S. state, including lawyers, accountants, auditors, teachers, nurses, engineers, psychologists, barbers and hairdressers. Occupational licensing directly affects 29 percent of U.S. workers, more than those affected by either minimum wage or unionization (Kleiner and Krueger 2010, 2013). Moreover, while the number of licensed occupations is rising, the proportion of the workforce being represented by trade unions is falling. Hence, an understanding of the determinants of licensing restrictions is growing increasingly important.

This paper explores the possibility of a link between the number of individuals attempting to enter a profession (potential supply) and the stringency of the entry requirements. While the existence of such a relationship is generally accepted in the literature (a summary is provided in Section 2), there is no direct evidence as to whether potential labor supply affects entry requirements. This may be due, in part, to the difficulty of measuring the stringency of entry requirements: while licensing boards may adjust the difficulty of the exams, their behavior is not generally observable to the researcher.

This work exploits an unusually rich panel data set on the U.S. market for lawyers. In this market, accurate data is available on bar exam difficulty, the number of exam candidates and exam outcomes. Detailed data on candidate ability can also be procured. Another factor making this market well suited for the present study is that the structure of the bar exam remains the same in the states and years in the sample, whereas the exam difficulty and the number of candidates vary significantly.

There are large discrepancies in exam difficulty across states. For example, when holding candidate ability constant, a change in exam difficulty from the standard in Alabama to the standard in California would imply a drop from 79% to 39% in the pass rate.[1] States with more numerous candidates tend to have more difficult examinations (holding candidate ability constant). Also using within-state variability, I find a positive correlation between the number of candidates and bar exam difficulty. Accounting for the possible endogeneity of potential supply increases the estimated correlation between the two variables. Overall, the paper shows that minimum entry requirements are relative standards, which are highly correlated with potential labor supply in the profession.

The magnitude of the estimated correlation suggests that licensing boards may significantly respond to changes in potential supply. Doubling the number of exam candidates is consistent with an increase of about 8 percent in exam difficulty. This implies that the actual increase in successful candidates may be about half of the increase that would have taken place without increases in standards. Thus, the licensing exam may partially shelter the market from supply shocks. More generally, it may affect the return to earning a professional degree and could dampen the impact of labor market policies targeted at increasing labor supply. Given the scale of public expenditure on education, it is important to understand whether professional licensing may influence the impact of such public investment on the labor supply.[2] Finally, professional licensing may also affect diversity in the profession. Since the service industry is a growing source of employment in developed economies, access to licensed professions may become an increasingly sensitive issue. This is particularly true for minorities, who provide a growing proportion of workers in less skilled licensed professions. The results of this paper are also relevant for the debate on the causes and consequences of occupational licensing and the applicability of

[1] I use for comparison a normal score distribution, with a mean equal to the mean bar exam score and the variance equal to the mean variance in the U.S. over the period 1981-2003. The grading procedures for the bar exam are described in Section 3.

[2] In 2007, OECD countries devoted 13 percent of total public expenditure to education, of which 3 percent to tertiary education (OECD 2010).

competition rules in professional markets in the U.S. and the European Union (Andrews 2002; Paterson, Fink and Ogus 2003; European Commission 2004).

2 Related literature

The stringency of entry requirements is the key variable controlled by licensing boards. The stated objective of entry examinations is uniquely to protect the public from unqualified practitioners. In fact, when standards are changed, there is typically no reference to changes in market conditions. However, there is agreement among economists that minimum standards are expected to vary depending on (potential) labor supply in the profession, since their impact on social welfare and salaries in the profession crucially depends on the availability of potential entrants. Independently of the exact objective function of licensing boards, then, potential supply is a key determinant of licensing stringency.[3] However, there is surprisingly little empirical research on the subject. In practice, little is known on how and why entry requirements change.

In one of the early contributions to the literature on licensing, Maurizi (1974) finds cross-sectional evidence of a negative correlation between the number of applicants and the pass rate on professional exams. He suggests that this correlation may be evidence of licensing boards increasing exam difficulty in response to excess supply. Although this evidence is intriguing (and similar results are obtained with my data, see Figure 1), using pass rates as a measure of licensing strictness has clear limitations, given that they

[3] There are two main views of licensing. According to Adam Smith (1776, I.x.c.5), the objective of licensing requirements "is to restrain the competition to a much smaller number than might otherwise be disposed to enter into the trade". According to this classic view, licensing is an inefficient institution that allows practitioners to capture monopoly rents by restricting entry (Friedman and Kuznets 1945, Stigler 1971). More recent theoretical studies have focused on the existence of asymmetric information on the quality of professionals (Akerlof 1970, Leland 1979, Shaked and Sutton 1981, Shapiro 1986). In the presence of asymmetric information, the licensing board takes into account both the quality-enhancing and competition-reducing effects of entry requirements. In this setting, if the objectives of the licensing board correspond to social welfare, licensing may be socially beneficial (the public interest theory of licensing, Leland 1979).

depend both on exam difficulty and candidate ability.

Leffler (1978) attempts to overcome this problem by developing a proxy for licensing difficulty in the market for physicians. Since candidates can take either a state or a national examination, the proportion of candidates choosing the state exam is used to develop a proxy for state exam difficulty. Although this is a significant step forward in measuring the stringency of entry requirements, the indirect procedure makes this proxy very imprecise. Moreover, candidate ability remains unobservable, and endogeneity may seriously affect the analysis (p.182).[4]

A related stream of literature has focused on the effect of licensing on wages and on the quality of professional services (Shepard 1978, HaasWilson 1986, Kleiner 1990, Kleiner and Kudrle 2000, Kugler and Sauer 2005, Timmons and Thornton 2008), and labor mobility (Pashigian 1979, 1980). Harrington and Krynski (2002), and Harrington (2007) study the impact of professional licensing in the funeral industry. Federman, Harrington and Krynski (2006) analyze the effect of state licensing regulations on low-skilled immigrants. Law and Kim (2004) study the historical origins of licensing, and Law and Marks (2009) the impact of licensing on minorities in the progressive era. Pagliero (2010) exploits changes in bar exam difficulty to estimate the effect of licensing requirements on entry-level salaries in the legal market. Winston, Crandall and Mahestri (2011) discuss the current policy debate on the regulation of the legal market.

All these studies focus on estimating the effects of licensing regulation on economic outcomes, implicitly assuming that licensing requirements are exogenously given. This paper departs from this stream of literature, as it does not focus on the effects of licensing regulation, but rather on the determinants of the stringency of entry regulation.[5]

[4] Kleiner (1990) provides a replication and time-series extension of the model first estimated by Maurizi (1974).

[5] Pagliero (2011) also looks at licensing standards as an endogenous outcome of regulation. However, the objective of the paper (identifying competing models of licensing) and the empirical strategy are different.

3 Brief overview of the bar exam and the data

The structure of the bar exam is the same in almost all states and has remained stable over the past two decades. The exam is administered twice a year, in February and July.[6] It consists of two components: the Multistate Bar Examination (henceforth MBE), a standardized test, and essay and case questions. The MBE contains 200 multiple choice questions developed by the National Conference of Bar Examiners, who are also responsible for correcting this portion of the exam. Using the results of a small sample of questions, which are repeated in different examinations over time and across states, scores are scaled so that any single MBE score represents a standard level of performance, regardless of when and where the exam is taken. Hence, the mean MBE score for candidates taking the exam in a given state and year is a cardinal measure of their average quality, and exam results can therefore be compared across states and years.[7]

Essay and case questions are set by state boards and graded at the state level, according to criteria set by each board.[8] In this case, a particular score does not necessarily correspond to a standard level of performance across states and years. However, states in my sample have introduced essay score scaling. The most common scaling procedure is mean and variance scaling. Mean and variance scaling requires that each essay score be transformed so that the mean and variance of the distribution of *scaled* essay scores is equal to the mean and variance of the standardized test scores (for each exam). The scaled essay scores are therefore not affected by exam-specific unobserved differences in exam difficulty or in the severity of grading procedures (Crocker and Algina 1986, Peterson et

[6]Exceptions are Delaware, Nevada and North Dakota, where the bar exam is held only once a year.

[7]A more detailed description of the MBE can be found at http://www.ncbex.org. A similar standardized test is the Graduate Record Examination (GRE), often used in the admission process to graduate courses.

[8]Some states have recently started to use essay and case questions developed by the National Conference of Bar Examiners (known as the Multistate Essay Examination and Multistate Professional Test). When this is the case, the Conference provides state boards with possible exam questions and some analysis of the issues involved in each question in order to facilitate grading. Even when using this service, state boards grade the answers independently, using locally-set standards.

al. 1993).[9]

Under the assumption that the unobserved candidates' quality is unidimensional, the *overall* scores (the weighted average of the standardized test and *scaled* essay test score) thus share the same metric across states and years and can be compared.[10] Since the pass-fail decision is based on overall scores, the observed minimum quality standards for each state also share a common metric and provide a simple measure of exam difficulty. (In the rest of the paper, I will refer to the overall minimum quality standard as exam difficulty, or the minimum standard).[11] Data on minimum standards is available from either 1984 or from the introduction of comparable standards (reported in Table 1, column 1), whichever is later, to 2012.[12] Table 1, column 2 reports any changes in minimum quality standards, while Column 3 reports the corresponding date of each change. Column 4 reports the minimum quality standard in the last year of the sample. Table 1 thus provides sufficient information for reconstructing the time series of the minimum standard in each state.

Minimum quality standard data is matched with the number of total and successful candidates for each examination, which is available from the National Conference of Bar Examiners for each state and year. The data set also includes data on MBE scores, consisting of MBE mean scores at the state level for each examination. Exam-specific

[9] An alternative scaling procedure is quantile by quantile equating. The results of the two techniques are not necessarily the same but differences are empirically small (see Lenel 1992).

[10] The assumption that the unobserved candidates' quality is unidimensional is consistent with licensing boards setting a single threshold for the overall scores, and not two separate thresholds for each component of the exam. Overall scores range between 0 and 200, like the MBE scores.

[11] The weights given to the two exam components may vary across states. Empirically, the weight given to the standardized test varies between 50 percent and 65 percent. For realistic distribution of scores and standards, however, these differences do not affect the comparability of minimum standards.

[12] The main source of standard and grading procedure data is The Comprehensive Guide to Bar Admission Requirements, published annually by the American Bar Association and the National Conference of Bar Examiners. This source is complemented by information from various issues of The Bar Examiner, published by the National Conference of Bar Examiners (NCBEX). When standards are comparable, but not expressed on a 0-200 point basis, the standards have been converted to a 0-200 basis to increase the consistency of Table 1. In the Comprehensive Guide there is some uncertainty as to when some standards changed. Wherever possible, additional sources have been used to pinpoint the exact date of change. In the few cases where no such data was available, the earliest date compatible with the information in the Comprehensive Guide was used.

information was furnished by the state Bar Association or the Supreme Court office responsible for administering the exam. I aggregate the information at the yearly level by summing the number of total and successful candidates for the two exams each year and calculating the mean MBE score (weighted by the number of candidates).

3.1 Data limitations

Data availability has been a long-standing issue in the literature on licensing.[13] In this context, my data set provides an exceptionally rich source of information on licensing board behavior. Still, there are three important limitations in the data set, summarized in Table 2. First, information on exam difficulty is not available for all the states (Table 1 reports information on minimum standards in the 37 states with observable exam difficulty). The reason for this is that the introduction of comparable standards based on standardized scores is a relatively new development in entry examinations. Some states still use grading procedures that do not allow quantitative comparison. Moreover, although they follow the grading procedures described above, some states do not publish their minimum standards. Table 3 provides summary statistics for the states with observable and unobservable exam difficulty. States with observable exam difficulty have very similar pass rates and number of candidates per capita, but larger populations and numbers of bar exam candidates.

The second limit to the data set concerns the availability of mean MBE score data (measuring bar exam candidates' quality), since only 12 licensing boards chose to provide this information. This problem is mitigated by using the available data on candidates' quality, pass rates and exam difficulty to construct an estimated measure of candidates' quality in each state and year. While closely matching the observed quality, when avail-

[13] Kleiner (2006, p.199) notes that "...perhaps the largest barrier standing in the way of analysis of occupational licensing is that there is no well-organized national data set waiting to be exploited. (...) Moreover, state licensing boards often are reluctant to provide (...) information to the researchers".

able, this estimated quality measure allows use of information for additional states and years in the empirical analysis. Appendix 1 describes in detail the construction of the data set. The third limitation of the data set is that exam difficulty is not frequently changed within a given state (see Table 1), as changes in admission rules typically involve a lengthy bureaucratic process. This limits the inferences one can make about state specific responses to changes in potential labor supply. Table 4 provides summary statistics for the states that changed exam difficulty. Relative to the states that did not change exam difficulty (Table 5), they have similar bar exam difficulty, candidates' quality, pass rate, and number of candidates per capita. However, they are larger in terms of state population and number of bar exam candidates.[14]

3.2 The Maurizi curve revisited

The negative correlation between pass rates and the number of exam candidates found by Maurizi (1974) is also a robust feature of my data (Figure 1).[15] Maurizi interpreted this negative correlation as evidence that licensing boards respond to increased potential supply by increasing entry requirements. Figure 2 challenges this interpretation by showing that candidate quality differs across states, and that states with better candidates tend to have higher pass rates. Clearly, pass rates are jointly determined by exam difficulty (chosen by licensing boards) and candidate quality (which is not controlled by licensing boards). Hence, studying exam difficulty, instead of pass rates, is a significant step forward in studying licensing board behavior.

Figure 3 shows that states with more candidates seem to have slightly more difficult

[14] One possible explanation for the difference between states with observable and unobservable exam difficulty is that the introduction of comparable standards involves some fixed costs, which can be shared among a larger number of applicants in larger states. Similarly, the different size of states that did and did not change exam difficulty could be explained by fixed costs in changing the minimum standards.

[15] In Figure 1, the number of candidates is divided by the state population, but similar patterns emerge without dividing by the state population, or dividing by the number of lawyers in the state.

examinations, although the correlation is not statistically significant.[16] Figure 4 shows that states with better candidates tend to have more difficult examinations. California, for example, has relatively difficult examinations and a large number of relatively good candidates. Alabama, on the other hand, has a small number of relatively weak candidates. Figures 3-4 describe the large amount of variability in exam difficulty, number and quality of candidates in my data (see also the summary statistics in Tables 4 and 5). The patterns described in these figures are interesting descriptive results, but do not necessarily capture causal relations. In particular, Figure 3 is very unlikely to describe the impact of potential labor supply on licensing exam difficulty. This is for two reasons. First, simple correlations may suffer from serious omitted-variable bias. In fact, the number and the quality of candidates are likely to jointly impact licensing board decisions.[17] Second, the variability in the number of bar exam candidates (and also their quality) is unlikely to be exogenous, since exam candidates may choose if and where to take the bar exam. These are the main topics of the next section.

4 Empirical results

I estimate regressions of the general form

$$D_{i,t} = b_0 + N_{i,t-1}b_1 + q_{i,t-1}b_2 + \lambda_t + \delta_i + u_{i,t} \tag{1}$$

where $D_{i,t}$ is the exam difficulty in state i and year t; $N_{i,t}$ is the log of the number of candidates; $q_{i,t}$ is the estimated average quality of candidates.[18] λ_t and δ_i are state and

[16] Similar results hold without dividing by the state population or dividing by the number of lawyers in the state.

[17] The two variables will generally be correlated in any sample, since they describe two features of the same group of individuals (who choose to take the licensing exam).

[18] Estimates are accurate, and using the observed MBE scores, when available, does not affect the results (see Section 3.1 and Appendix 1 for details).

year fixed effects, and $u_{i,t}$ is the idiosyncratic error term. Since changes in minimum standards are determined and published in advance, my empirical specification (1) allows for a lagged effect of potential supply on exam difficulty.[19]

Cross Sectional Evidence

Table 6, columns 1-2 report OLS estimates of the pooled and cross sectional correlation between potential supply and difficulty of the bar exam. The estimated coefficients of the number and the quality of candidates are positive and significantly different from zero. Adding year fixed effects in column 2 does not significantly affect the coefficients. A ten percent increase in the number of candidates implies an increase of about 0.08 in exam difficulty (measured on the MBE scale). States with more candidates tend to have more difficult entry examinations.[20] An increase of one in candidates' average quality implies an increase of about 0.7 in exam difficulty.

Before commenting on the magnitude of the estimated coefficients, it is important to note that there are three reasons why the results in columns 1-2 may be a biased estimate of the impact of potential supply on exam difficulty. First, bar exam candidates are likely to be aware of the systematic differences in exam difficulty across states, so relatively good candidates may be more likely to take the bar exam in states with higher standards. Weaker potential candidates may be more likely to move to a different state or a different type of job. Such a selection mechanism may lead to an upward bias in the estimated impact of quality, and a downward bias in the impact of the number of candidates. Second, candidates preparing for a difficult exam may study more than those taking an easier examination. Reverse causality may thus further bias upwards the impact of candidate quality. Third, there may be some omitted state characteristics that independently lead to high entry standards and high potential supply. For example, a

[19] Section 4.1 explores alternative specifications of (1).

[20] Similar results are obtained dividing the number of candidates by the population of the state (see Table A1, columns 1 and 2).

state may have a particularly good educational system, thus leading to a large supply of bar exam candidates and high average quality. In principle, such a culture of excellence may also affect bar examiners, who may then set relatively high standards for admission to the profession.[21]

Within State Evidence

Table 6, column 3 reports the results controlling for state-specific fixed effects, so that state-specific unobserved characteristics can be accounted for. The coefficients of the number and quality of bar exam candidates are still positive. The magnitude of the coefficients is quite different from that obtained using cross sectional variability. The coefficient of the number of candidates is twice that obtained in columns 1-2, while the coefficient of quality is about half. These changes are consistent with a significant self-selection of candidates across states, and/or exam difficulty affecting candidates' effort in exam preparation, and/or unobserved state characteristics affecting both potential supply and exam difficulty. Including both state and year fixed effects in column 4 does not affect the estimated coefficients. This suggests that year fixed effects are not capturing the effect of important omitted variables correlated with the regressors and the error term in (1).

I next focus on the sub-sample of states that changed their entry standards. This is because unchanged standards over a long period may derive from unobserved differences in how licensing boards choose entry standards, or particularly large adjustment costs. This implies a substantial reduction in sample size (from 37 to 14 states), but also a reduction in unobserved heterogeneity. For states that changed exam difficulty, the coefficient of the number and quality of bar exam candidates is still positive and significantly different from zero, and the magnitude of the effects is larger than before. [22]

[21] Similarly, a state may have a particularly high demand for high quality lawyers, thus leading to a large supply of high quality candidates. This pattern of specialization may also affect the characteristics of bar examiners, leading to a spurious correlation between potential supply and exam difficulty.

[22] This increase is not surprising. The estimated coefficients cannot be smaller when using only the data for states that changed difficulty. In fact, when using all the states, the estimated coefficients are an average of the coefficients for states that changed exam difficulty and states that did not, but the

Consider the effect of doubling the number of candidates (by drawing from a given score distribution) in a hypothetical state with mean exam difficulty and mean score distribution: with no change in exam difficulty, about 73 percent (which is the average pass rate) of the new candidates would pass the exam. In practice, however, the results suggest that the exam difficulty will grow in response to the increased potential supply by about 3.5 (see Table 6, column 8), which implies that only 61 percent of the new candidates will pass the exam.[23] Hence, about 15 percent of the effect on the number of successful candidates at the entry examination is offset by the increase in exam difficulty. The magnitude of the offsetting effect of exam difficulty is significant even if we use the smaller estimated coefficients in Table 6, column 4. Consider now an increase in candidates' quality of 5 on the MBE scale (approximately the difference between mean candidates' quality in Maryland and Virginia in Figure 2). This implies an increase in the pass rate from 73 to 84 percent. However, because of the increase in exam difficulty implied by the coefficient of candidates' quality (Table 6, column 8), only 78 percent of the new candidates will pass the exam. Hence, the estimated coefficients are consistent with about half of the effect on the pass rate being offset by changes in exam difficulty.

Overall, the state fixed effects account for significant sources of potential bias in the estimates, but endogeneity may still be a problem when using within state variability: increases over time in exam difficulty may in fact result in weaker candidates taking the exam in a different state, moving to a different profession, or studying harder for the exam. All of these would imply a downward bias in the coefficient of the number of candidates,

---

latter are zero, since there is no within-state variation in the dependent variable.

[23] These figures are obtained for an hypothetical state with mean exam difficulty (134) and Gaussian score distribution with mean equal to the mean MBE score (141) and standard deviation equal to 12 (the mean standard deviation). The relation between pass rate and exam difficulty is non linear. Consider n candidates taking a professional examination. Each candidate receives an overall exam score s, which is a random draw from the distribution $F(s)$. The difficulty of the exam, D, is the minimum quality allowed in the market. All candidates who score at or above the threshold pass the exam and enter the profession. The number of successful candidates is then $L = [1 - F(D)]n$ and the pass rate $L/n = 1 - F(D)$ is non linear in D.

and an upward bias in the coefficient of quality. Hence, the offsetting effect of increases in the number of candidates is likely to be larger than in Table 6, while the offsetting effect of changes in quality smaller. This issue can be addressed using instrumental variables; these must be correlated with the number and quality of candidates but not with the error term in (1).[24]

Instrumental Variables

Bar exam candidates typically need to have a law school degree, which usually requires at least three years of law school. Within any given state, then, the number of students admitted to law school three years before a given bar exam is likely to be correlated with the number of bar exam candidates (after controlling for state fixed effects), since many students take the bar exam in the state where they studied. A similar argument can be made for the quality of first-year law school students, which is likely to be correlated with the quality of bar exam candidates three years later (after controlling for state fixed effects).[25] The magnitude and statistical significance of these correlations can be estimated; hence the empirical relevance of the instruments can be tested.

Detailed information on class size for each law school in each year is available from the American Bar Association Official Guide to Law Schools. Information on the 25^{th} and 75^{th} percentiles of the LSAT score distribution of admitted students for each law school (from the same source) provides a good measure of the candidates' average quality (the LSAT is a standardized examination used by law schools to screen applicants). The two instrumental variables are then the number of matriculated students and the LSAT mean score in law schools for each state and year.[26] Since there is no law school in Alaska, the

[24] A second reason for using instrumental variables is to account for possible measurement error in q_{it}.

[25] Such correlation is expected to be positive, but not perfect, as students from some top schools (e.g., Harvard) tend to allocate themselves nationally. However, Oyer and Schaefer (2010) provide evidence that law firms tend to hire largely from local law schools.

[26] As the LSAT is a standardized examination, results can be compared across states and years. The mean LSAT score for each school and year is estimated as the average of the 25th and 75th percentiles. This will be a biased estimate of the mean LSAT score if the LSAT score distribution is not symmetric. However, unbiasedness (or even consistency) of this estimator is not a requirement for using the estimated

corresponding observations must be dropped when using instrumental variables.[27]

Class size and quality of students admitted to law school are unlikely to be affected by changes in exam difficulty three years later. Changes in standards are not typically announced or even decided so far in advance. The IV approach assumes that the number and quality of students enrolling in law schools in year t are not correlated with the unobserved determinants of exam difficulty ($u_{i,t+3}$) in year $t + 3$. Law school students can still base their choice on any observable or unobservable characteristic of the state, captured by the state fixed effects. For example, students may choose where to attend law school based on the specificities of the demand for legal services, the characteristics of law firms, and of the matching process of lawyers and law firms in each state.

A second requirement for the instrumental variable approach is that the instruments can be reasonably excluded from equation (1). This is realistic, since the number and quality of first-year law school students matters for outcomes in the legal market only to the extent that they take the bar exam three years later. The behavior of professional associations is not further restricted. Professional associations can still set exam difficulty as a function of state-specific characteristics of the supply and demand for lawyers in the state (captured by the state fixed effects).

Table 7, columns 1-2 report the 2SLS estimates of the coefficients of potential supply on exam difficulty, controlling for state fixed effects. Column 1 includes all the states, column 2 excludes the states with no change in exam difficulty. Table 8, columns 1-4 report the first stage estimates. As expected, column 1 shows that the number of

mean LSAT score as an instrument for the 2SLS estimation of model (1). Both LSAT score and class sizes are aggregated at the state-year level (average LSAT scores are appropriately weighted).

[27]Data on the instruments is available from 2010 to 1993, then there is a 7 year gap between 1992 and 1986, and finally one usable set of observations for 1985. Given the large gap in the instruments, observations for the instrumental variables before 1993 were dropped. This implies that the number of observations in Table 7 is significantly smaller than for the OLS regressions (given the lags in the IVs, we can only use observations between 1997 and 2010, rather than 1989-2010). However, when the values between 1986 and 1992 are interpolated, the number of observations can be significantly increased. The results are not affected (see Table A1, columns 7-8).

matriculated students is positively correlated with the number of bar exam candidates three years later, after controlling for state fixed effects. Similarly, column 2 shows that the quality of law school students (LSAT score) is positively correlated with the quality of bar exam candidates three years later (MBE score), after controlling for state fixed effects. These estimated effects are statistically different from zero at a one percent confidence level. The partial R^2 of the first stage regressions show that the instruments account for a significant portion of the variability of the two endogenous regressors. The F-tests of the excluded instruments in the first stage are large.[28]

The 2SLS estimates in Table 7, columns 1-2 show that the coefficient of the number of bar exam candidates on exam difficulty is again positive and significantly different from zero at conventional levels. Its magnitude is larger than in Table 6, columns 3 and 7, which report the corresponding OLS estimates. In order to evaluate the magnitude of the coefficients, consider again the effect of doubling the number of candidates in a hypothetical state with mean exam difficulty and mean score distribution. The 2SLS results suggest that the exam difficulty may grow by about 11 (Table 7, column 2), which implies a pass rate of 36 percent. Hence, about half of the effect on the number of successful candidates may be offset by the increase in exam difficulty. The increase in the magnitude of the coefficient of the number of candidates is consistent with some remaining endogeneity in the within state OLS regression. In fact, selection effects may lead to a negative correlation between exam difficulty and the number of bar exam candidates thus biasing downward the impact of the number of candidates in OLS regressions.

The 2SLS coefficient of candidates' quality is smaller than in Table 6 and not significantly different from zero. This also supports the idea that within state results were affected by some remaining endogeneity and biased upward. The IV estimate is also less precise than the corresponding OLS estimate in Table 6. Hence, on the one hand, the

[28]The null of weak instruments can be rejected using the thresholds of Stock and Jogo (2002). The weak instrument problem is generally more relevant in overidentified than in just identified models.

null of no impact of candidates' quality cannot be rejected. On the other, the power of the test is rather small, and the 95 percent confidence intervals still include values which imply a significant effect of candidate quality on the number of entrants in the profession.

The 2SLS results are robust when both state and year fixed effects are included (Table 7, columns 3-4). However, the first stage coefficients in Table 8, columns 5-8 are not precisely estimated because of the reduction in the degrees of freedom, and also the partial R^2 are smaller. Hence, 2SLS results with state and year fixed effects should be interpreted with caution. However, given the little effect of adding year fixed effects in Table 6, columns 2 and 6, there seems to be no evidence suggesting that year fixed effects capture important sources of potential endogeneity.

4.1 Additional results

Admission on Motion and Additional Requirements

Although the bar exam is by far the most common admission procedure, admission to the bar is sometimes possible without taking the bar exam. Lawyers licensed in other states may be admitted on motion, typically on the basis of reciprocity agreements.[29] In principle, the number of these admissions may affect lawyer supply and may therefore impact exam difficulty. When I include the number of admissions on motion in my specification, the results are not substantially affected (Table A1, columns 3-6). The number of lawyers admitted on motion seems to have a positive impact on exam difficulty. Although the magnitude of such an effect is small, this finding reinforces the conclusion that bar exam difficulty is affected by the supply side of the market for lawyers.

In addition to the bar exam, admission to the legal profession also requires passing

[29] In 2004, there were fewer than 6,000 admissions on motion, but over 49,000 by examination (NCBEX). Only in D.C. is this the main mode of admission to the bar. However, the results are not affected when I exclude the corresponding observations from the sample. Direct admission of law school graduates (by diploma privilege) is significant only in Wisconsin, which has been excluded from the sample.

the Multistate Professional Responsibility Examination (MPRE). The MPRE is a standardized multiple-choice examination based on law governing the conduct of lawyers, including the disciplinary rules and principles of professional conduct. Data on the minimum scores required for passing this exam is available for each state.[30] If higher minimum MPRE scores were used by licensing boards as a substitute for a more difficult bar exam, one would expect that states with higher MPRE thresholds would have an easier bar exam, all other variables being held constant. However, there are virtually no changes in MPRE minimum scores, so the regression results including state fixed effects cannot be affected.[31]

Admission to the bar also depends on meeting educational standards and moral character and fitness requirements. Although educational standards impose significant costs on candidates, they do not vary significantly across states, and are fixed for a given state.[32] The procedures for moral character and fitness evaluation cannot easily be compared across states and years, but there is no evidence of significant within-state variability.[33] Moreover, the existing evidence suggests that these procedures directly affect only a very small number of candidates.

Alternative specifications, GDP per capita, and state-specific time trends

Table A1, columns 9-10 reproduce the results in Table 7, columns 1-2 excluding the observations after the fourth year following each change in standards. When more than one change in exam difficulty occurs in a given state, all the observations after the fourth

[30] The MPRE is required for admission to the bar in all but three US jurisdictions. Data on the number of MPRE candidates and their performance is not available. As opposed to the bar exam, most candidates take the MPRE during their second or third year of law school, well before applying for the bar exam. Overall, the complexity of the subjects tested on the MPRE is much lower than on the bar exam. The breadth of the subject also suggests that much less effort is usually required to pass (this is consistent with the lower emphasis given by review courses to MPRE preparation).

[31] When we exclude the one state, Missouri, in which a change in MPRE standard occurred in the sample period, there is no change in the results.

[32] Overall, most bar exam candidates hold a law school degree. In 2004, for example, only two states had more than 10 percent of total candidates without a US law school degree.

[33] In 2005, 19 states have no published character and fitness standards in The Comprehensive Guide to Bar Admission Requirements.

year following the first change are dropped. The objective is controlling for possible long-run responses of bar exam candidates to changes in difficulty. In fact, excluding these observations implies that the number of first year law school students and their quality (the excluded instruments) cannot possibly depend on the observed changes in exam difficulty (after controlling for state fixed effects). The results are not substantially affected.

Equation (1) assumes that licensing boards react to a single lag in potential supply. However, licensing boards may consider changes in potential supply over multiple years when revising entry standards. Table A2 provides a second set of robustness results describing the effect of the average potential supply in previous years. The coefficient of potential supply is again positive and statistically significant.[34]

The main results presented in Tables 6 and 7 control for time-invariant state characteristics and time effects that are common to all the states. These regressions do not control for time-varying state-specific variables, which may capture long term changes in the local labour markets. Columns 1-4 in Table A3 report OLS estimates of equation (1) including real GDP per capita in each state as a control variable. Columns 5-8 include the manufacturing share of GDP in each state. Finally, columns 9-10 include state-specific time trends. Table A4 reports the corresponding 2SLS estimates including real GDP per capita and the share of manufacturing in each state. The coefficients of the number of candidates are not significantly affected both in terms of magnitude and statistical significance. This suggests that the correlations described in the previous tables were not due to state-specific economic growth, or changes in the sectorial composition of the economy.[35]

[34] For 2SLS estimation (columns 3 and 6), the definition of the instruments is changed accordingly. The results are not affected using weighted averages with decreasing weights for longer lags. Similar results are obtained directly including lags of the number and quality of exam candidates (although standard errors increase).

[35] While the results cannot be affected by time invariant characteristics of the licensing boards, one cannot rule out that changes in how the bar associations are organized (or in the influence of pro-lawyer

5 Conclusions

Occupational licensing is one of the most important labor market institutions, yet the actual behavior of licensing boards is rarely examined because of a lack of data and the complexity of licensing requirements. While the existing literature typically assumes that licensing regulations are exogenous, this paper assumes that regulation is endogenous and studies the possible effect of potential supply on the difficulty of the entry examination. The main finding is that potential supply and the difficulty of the bar exam are correlated, even after accounting for a number of confounding effects. This is consistent with supply considerations being an important determinant of licensing requirements. In particular, increases in the number of potential entrants seem to lead to more stringent entry requirements, even after controlling for the average quality of candidates.

The existence of a correlation between potential supply and entry requirements has a number of potentially important implications. Professional markets may be sheltered from the impact of policies increasing potential supply. For example, a policy increasing the availability of education, while holding quality constant, may increase the number of candidates but only partially increase labor supply. A second implication is that licensing may affect how groups of different average ability are represented within a profession. Consider an increase in the number of exam candidates (as the outcome of a policy subsidizing education, for example) which equally affects all groups of candidates. As standards increase, groups with lower average performance become less represented among the successful candidates. For example, since bar exam outcomes vary dramatically across ethnic groups, this effect may significantly affect ethnic diversity within the legal profession (Wightman 1998, p.27, reports a 30 percent difference in pass rates be-

lobby groups) may affect the difficulty of the exam. Unfortunately, no systematic evidence is available on changes in these variables. In the market for physicians (also a licensed profession), Law and Hansen (2010) investigate how licensing board characteristics affect the frequency with which boards discipline physicians.

tween African American and white candidates). Even if a similar policy only increases the number of candidates from a group with relatively low average exam performance, a decrease in the fraction of successful candidates from such group cannot be ruled out.

From this point of view, it is worth considering the possibility of certification as an alternative to licensing (Kleiner 2006). Certification permits any person to practice the profession, but a public or private institution certifies those who have achieved a given level of competence. Certification may increase the effectiveness of policies affecting potential labor supply (without affecting the representation of minorities in the profession), while still revealing information on practitioners' quality.

At a more general level, this paper shows that entry requirements may significantly respond to market forces and, hence, are not uniquely based on some abstract notion of consumer protection, unrelated to market conditions.[36] The results of this paper are clearly consistent with licensing boards restricting entry into the profession in order to increase salaries (Friedman and Kuznets 1945, Stigler 1971). However, at least in principle, they could also be explained by public interest theory (Leland 1979), provided there were large asymmetric information and consumers' valuation for quality (as measured by the bar exam) were sufficiently high. In such circumstances, the social planner trades off the welfare gains from higher quality with those deriving from increased availability of professionals (and lower prices). Hence, a social planner could partly offset increases in potential supply by increasing exam difficulty.

In practice, it seems difficult to explain the results of this paper *entirely* based on the public interest theory. First, there seems to be no agreement in the legal profession on the extent to which performance at the bar exam is an appropriate variable to be used for screening potential entrants in the legal market (Society of American Law Teachers 2002). There is also little evidence that consumers value increases in exam difficulty (Pagliero

[36] The stated objective of the bar exam is "to protect the public, not to limit the number of lawyers admitted to practice" (ABA and NCBEX 2000, p. ix).

2011), which is a necessary condition for the explanation based on public interest theory. Moreover, there is evidence that increases in exam difficulty lead to increases in entry salaries in the legal profession, but no evidence that they disproportionately increase the lower deciles of the salary distribution, which would naturally occur if the bar exam screened out the worse potential entrants (Pagliero 2010). Finally, there is no evidence suggesting that differences in asymmetric information across states are large enough to justify the large differences in exam difficulty observed in the data. Still, one cannot rule out that asymmetric information could be more relevant in states with difficult exams (e.g., California) then in states with easier exams (e.g., Alabama).

In conclusion, the existence of a link between potential supply and the stringency of licensing regulation opens the debate on the impact of supply shocks and the effectiveness of public policy in professional markets. It also contributes to the ongoing debate on how the interests of the profession may affect labor market regulation.

References

Akerlof, G. (1970). The market for lemons: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics 89*, 488–500.

American Bar Association and National Conference of Bar Examiners (2000). *Comprehensive Guide to Bar Admission Requirements.* ABA and NCBEX.

Andrews, P. (2002). Self regulation by professions - the approach under EU and US competition rules. *European Competition Law Review 23*(6), 281–285.

Crocker, L. and J. Algina (1986). *Introduction to classic and modern test theory.* New York: Holt Reinhard and Winston.

European Commission (2004). *Report on Competition in Professional Services.*

Federman, M., D. Harrington, and K. Krynski (2006, May). The Impact of State Licensing Regulations on Low-Skilled Immigrants: The Case of Vietnamese Manicurists. *American Economic Review 96*(2), 237–241.

Friedman, M. and S. Kuznets (1945). *Income from Independent Professional Practice.* NBER.

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

Harrington, D. and K. Krynski (2002). The Effect of State Funeral Regulations on Cremation Rates: Testing for Demand Inducement in Funeral Markets. *Journal of Law and Economics 45*(1), 199–225.

Harrington, D. E. (2007, Fall). Preserving Funeral Markets with Ready-to-Embalm Laws. *Journal of Economic Perspectives 21*(4), 201–216.

Kleiner, M. (1990). Are there economic rents for more restrictive occupational licensing practices? Industrial Relations Research Association.

Kleiner, M. and A. Krueger (2010). The prevalence and effects of occupational licensing. *The British Journal of Industrial Relations 48*(4), 676–687.

Kleiner, M. and A. Krueger (2013). Analyzing the Extent and Influence of Occupational Licensing on the Labor Market. *Journal of Labor Economics* (forthcoming).

Kleiner, M. M. (2006). *Licensing Occupations: Ensuring Quality or Restricting Competition?* Upjohn Institute.

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

Kugler, A. and R. Sauer (2005). Doctors without Borders: The Returns to an Occupational License for Soviet Immigrant Physicians in Israel. *Journal of Labor Economics 23*(3), 437–465.

Law, M. and M. Marks (2009). Effects of occupational licensing laws on minorities: Evidence from the progressive era. *The Journal of Law and Economics 52*(2), 351–366.

Law, M. T. and Z. K. Hansen (2010). Medical licensing board characteristics and physician discipline: An empirical analysis. *Journal of Health Politics, Policy and Law 35*(1), 63–93.

Law, M. T. and S. Kim (2005). Specialization and regulation: The rise of professionals and the emergence of occupational licensing regulation. *Journal of Economic History 65*(4), 723–756.

Leffler, K. B. (1978). Physician licensure: Competition and monopoly in American medicine. *Journal of Law and Economics 21*(1), 165–186.

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

Lenel, J. (1992, May). Issues in equating and combining MBE and essay scores. *The Bar Examiner 61*(2), 6–20.

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

Oyer, P. and S. Schaefer (2010). Firm/Empoyee Matching: An Industry Study of American Lawyers.

Pagliero, M. (2010). Licensing Exam Difficulty and Entry Salaries in the US Market for Lawyers. *British Journal of Industrial Relations 48*(4), 726–739.

Pagliero, M. (2011). What is the Objective of Professional Licensing? Evidence from the US Market for Lawyers. *International Journal of Industrial Organization 29*(4), 473–483.

Pashigian, P. B. (1979). Occupational licensing and the interstate mobility of professionals. *Journal of Law and Economics 22*(1), 1–25.

Pashigian, P. B. (1980). Has occupational licensing reduced geographic mobility and raised earnings? In S. Rottenberg (Ed.), *Occupational Licensure and Regulation*, pp. 299–333. Washington, D.C.: American Enterprise Institute.

Paterson, I., M. Fink, and A. Ogus (2003). *Economic Impact of Regulation in the Field of Liberal Professions in Different Member States.* Institute for Advanced Studies, Vienna.

Petersen, N. S., M. J. Kolen, and H. Hoover (1993). Scaling, norming and equating. In Linn (Ed.), *Educational Measurement.* American Council on Education / Macmillan.

Shaked, A. and J. Sutton (1981). The self-regulating profession. *The Review of Economic Studies 48*(2), 217–234.

Shapiro, C. (1986). Investment, moral hazard and occupational licensing. *The Review of Economic Studies 53*(5), 843–862.

Shepard, L. (1978). Licensing restrictions and the cost of dental care. *Journal of Law and Economics 21*(1), 187–201.

Smith, A. (1776). *An Inquiry Into the Nature and Causes of the Wealth of Nations.* London: Methuen and Co., Ltd.

Society of American Law Teachers (2002). Statement on the bar exam. *Journal of Legal Education 52*(3), 446–452.

Stigler, G. (1971). The theory of economic regulation. *Bell Journal of Economics 2*, 3–21.

Stock, J. and M. Yogo (2005). Testing for weak instruments in linear iv regression. In

D. Andrews and J. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothemberg*, pp. 80–108. Cambridge: Cambridge University Press.

Timmons, E. J. and R. J. Thornton (2008). The effects of licensing on the wages of radiologic technologists. *Journal of Labor Research* (29), 333–346.

Winston, C., R. W. Crandall, and V. Mahestri (2011). *First thing we do, let's deregulate all the lawyers*. Washington, D.C.: Brookings Institution Press.

Appendix 1. Data construction and sources

The lack of complete information on MBE mean scores is the main limitation of the data set. However, one can infer MBE scores using information on pass rates and exam difficulty. The procedure is based on the assumption that the shape of the score distribution is stable over time. More specifically, I assume that the state specific score distribution is Gaussian, with a constant standard deviation.[37] Consider, for example, the exam outcomes in subsequent years for a given state, with a given *constant* exam difficulty. If the standard deviation of the distribution does not change over time, then changes in pass rate from one year to the next will reflect changes in the location of the score distribution. Since there is a one to one relation between pass rate and the mean of the score distribution, changes in pass rate can then be used to infer changes in mean MBE scores (details provided below).

There are advantages and disadvantages to taking this approach. The main advantage is that I can partly compensate for the difficulty in obtaining information from licensing boards. Second, I can efficiently use the information on the number of successful and unsuccessful candidates, which is very precise, complete and publicly available. Hence, I can increase the number of observations and significantly improve the quality of the inference.[38] The main drawback of this approach is that some assumptions about the score distribution are necessary. However, I test the robustness of the results using alternative assumptions and the results are not affected. Moreover, the proposed method seems reasonably accurate in capturing mean MBE scores variability, independently of the specific functional form assumptions. Overall, the advantages from efficiently using the available information on pass rates seem to outweigh the disadvantages from maintaining

[37] Although there is no direct evidence at the state level, the available information at the US level suggest that the overall shape of the score distribution is stable over time. Moreover, the standard deviation of the score distribution does not show any significant trend, suggesting that the assumption of constant standard deviation is realistic.

[38] The results are consistent with what I would obtain using the original data, but they are much more precise.

some assumptions on the score distribution.

The procedure for recovering MBE mean scores is the following:

1. Consider how exam-specific results are generated. Each candidate passes the bar exam if his/her overall score is above a given threshold. In exam k, in state i, candidates' scores are independent draws from a normal distribution with mean $\mu_{i,k}$ (equal to the mean MBE score for exam k) and unknown variance σ_i^2. Hence, the number of candidates passing the bar exam is defined, for each examination, by

$$P_{i,k} \equiv \left[1 - F\left(\frac{D_{i,k} - \mu_{i,k}}{\sigma_i}\right)\right] N_{i,k}. \tag{2}$$

   where $F(.)$ is the cumulative distribution function of the standardized normal distribution. The likelihood of observing P_k successful candidates out of N_k exam candidates is then

$$L = \prod_k F\left(\frac{D_{i,k} - \mu_{i,k}}{\sigma_i}\right)^{(N_{i,k} - P_{i,k})} \left[1 - F\left(\frac{D_{i,k} - \mu_{i,k}}{\sigma_i}\right)\right]^{P_{i,k}} \tag{3}$$

   where $D_{i,k}$ is the observed exam difficulty. Using information for the 12 states for which I can observe $D_{i,k}$, $\mu_{i,k}$, $N_{i,k}$, and $P_{i,k}$, maximization of the log-likelihood $\ln L$ provides estimates of σ_i.[39]

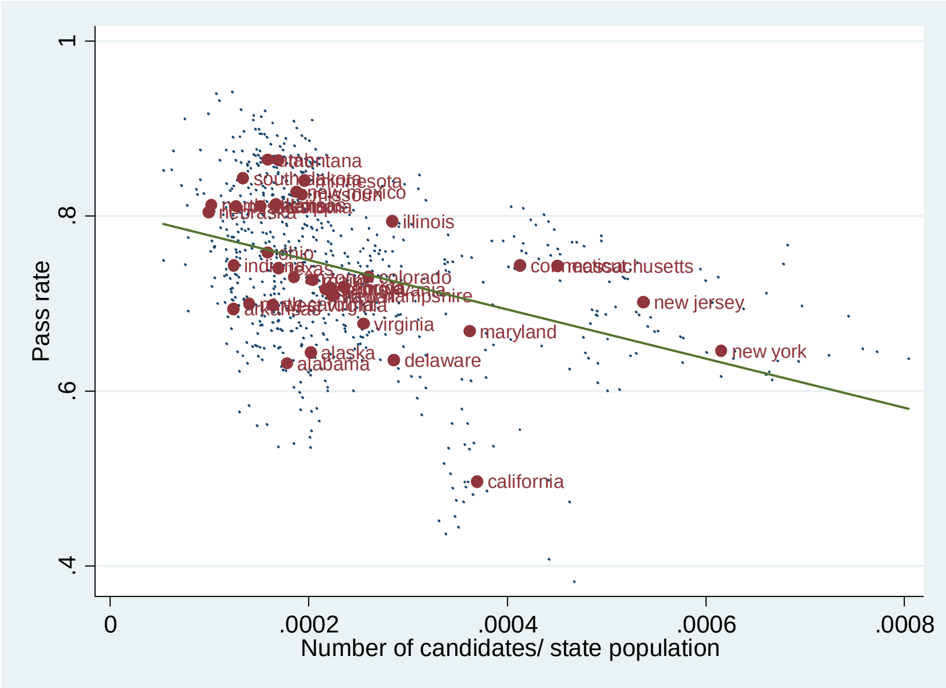
2. Having estimates of the standard deviation of the score distribution, one can recover the mean score even when it is not made available by licensing boards. Given how the pass rate is determined, and the fact that we exactly observe $D_{i,k}$, $P_{i,k}$, and $N_{i,k}$, $q_{i,t}$ can be computed inverting the identity (2),

$$q_{i,k} \equiv D_{i,k} - \widehat{\sigma}_i F^{-1}\left(1 - \frac{P_{i,k}}{N_{i,k}}\right) \tag{4}$$

[39] The results are not significantly different when I assume that scores have a beta distribution, which can accommodate some skewness in the score distribution.

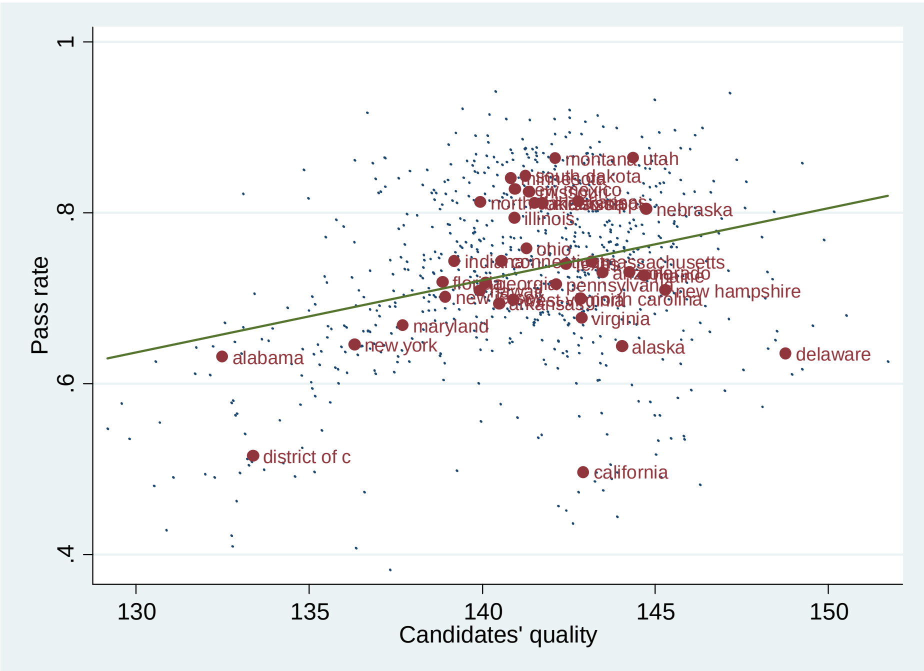
where $\widehat{\sigma}_i$ is the estimated standard deviation of the score distribution in state i (or the mean $\overline{\sigma} = \frac{1}{12}\sum_{i=1}^{12}\widehat{\sigma}_i$, if $\widehat{\sigma}_i$ cannot be estimated because $\mu_{i,k}$ is not available).

Figure 1 The Maurizi curve.



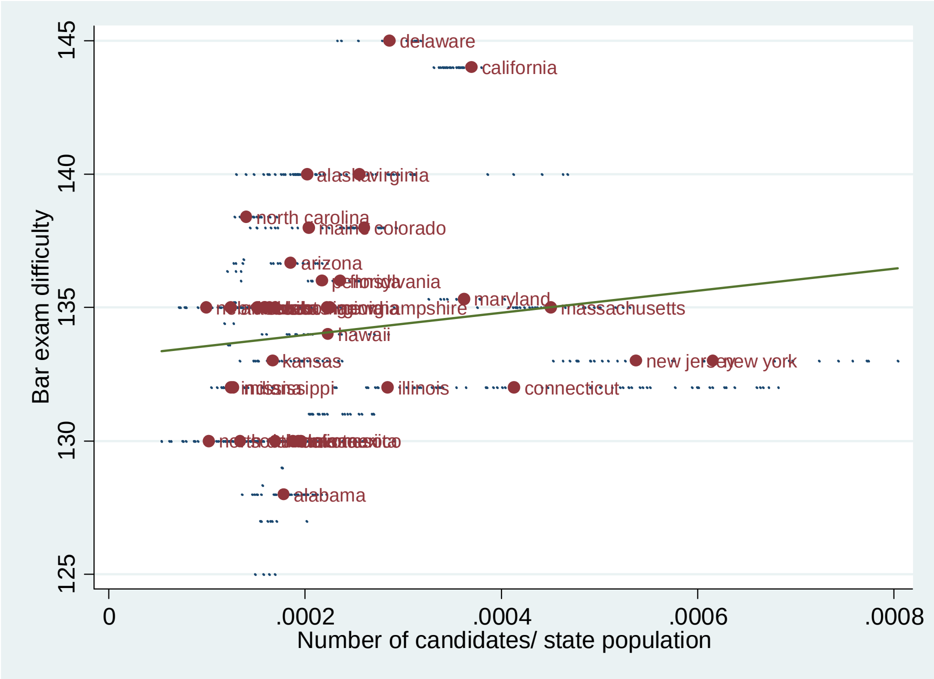
Note: D.C. is omitted from the graph as it is an outlier with a large number of candidates (0.0085) and low pass rates (0.52). The linear correlation is statistically significant at the 1 percent confidence level (and even more pronounced when including D.C.). The blue (small) dots describe state-year observations. The red dots correspond to state-specific mean values.

Figure 2. Pass rate and candidates' quality.



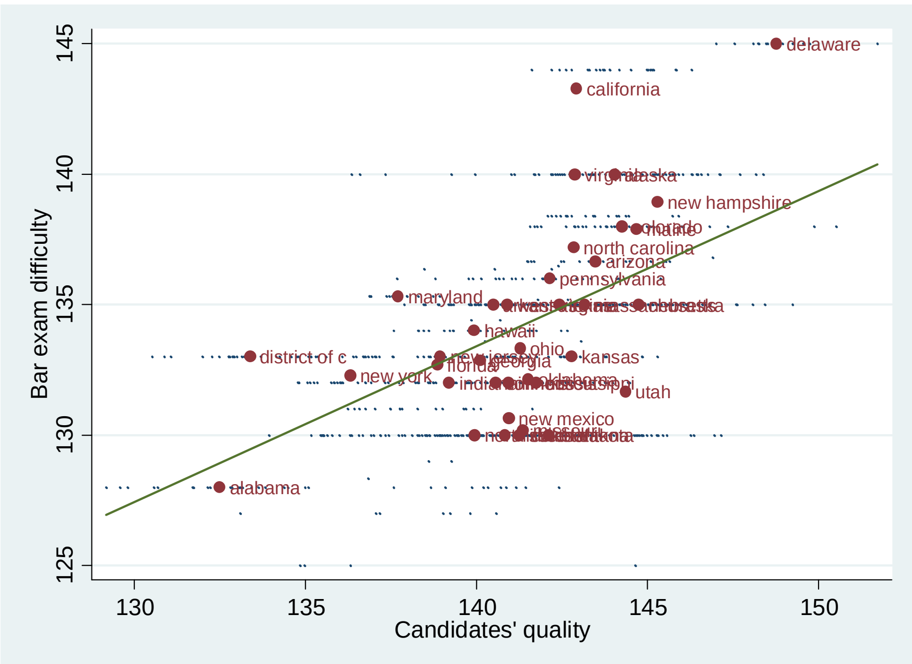
Note: The linear correlation is statistically significant at the 5 percent confidence level. The blue (small) dots describe state-year observations. The red dots correspond to state-specific mean values.

Figure 3. Bar exam difficulty and number of bar exam candidates.



Note: The linear correlation is not statistically significant. D.C. is omitted as it is an outlier with a large average number of candidates (0.0085). The correlation is unchanged when including DC. The blue (small) dots describe state-year observations. The red dots correspond to state-specific mean values.

Figure 4. Bar exam difficulty and candidates' quality.



Note: The linear correlation is statistically significant at the 1 percent confidence level. The blue (small) dots describe state-year observations. The red dots correspond to state-specific mean values.

Table 1. Bar exam difficulty.

State	Starting date of comparable standards	Observed changes in bar exam difficulty	Date of change	Bar exam difficulty in 2012 (0-200)
	(1)	(2)	(3)	(4)
Alabama	1990	-	-	128
Minnesota	1984	-	-	130
Missouri	1984	5, -3	1996, 2005	130
Montana	1999	-	-	130
New Mexico	1984	3, -3	1990, 96	130
North Dakota	1986	-	-	130
South Dakota	1989	-	-	130
Connecticut	1984	-	-	132
Illinois	2000	-	-	132
Indiana	2001	-	-	132
Mississippi	1995	-	-	132
D.C.	1984	-	-	133
Kansas	2000	-	-	133
New Jersey	1992	-2	1993	133
New York	1984	1	2005	133
Hawaii	1993	-	-	134
Arkansas	2002	-	-	135
Georgia	1984	5	1997	135
Massachusetts	1984	-	-	135
Nebraska	1996	-	-	135
New Hampshire	1984	-5	2007	135
Ohio	1984	-10, 3.33, 6.67	1992, 96, 97	135
Oklahoma	1984	2, 1, 4, 1	1991, 92, 95, 97	135
Texas	1994	-	-	135
Utah	1991	5	2006	135
West Virginia	1994	-	-	135
Maryland	2000	-	-	135.33
Florida	1984	2, 3	2003, 04	136
Pennsylvania	2001	-	-	136
Arizona	1991	-0.17	2012	136.5
Colorado	1987	-	-	138
Maine	1984	1, 2, 2, -2	1990, 92, 95, 2003	138
North Carolina	1984	-2.8, 0.8, 0.8, 0.8, 0.8, 1.6	1988, 90, 92, 94, 95, 96	138.4
Alaska	1992	-	-	140
Virginia	1998	-	-	140
California	1984	4	1990	144
Delaware	2000	-	-	145

Note: Bar exam difficulty is the minimum overall score required to pass the bar exam (measured on a 0-200 scale) in each state. Data on difficulty is available from either 1984 or the introduction of comparable standards (reported in column 1), whichever is later, to 2012. Column 2 reports changes in difficulty, while column 3 reports the corresponding date of each change. Column 4 reports difficulty in 2012. The information in Table 1 allows reconstruction of the time series of exam difficulty in each state.

Table 2. Overview of data availability on key variables.

	Observed?	Data availability	Used in the analysis?
Bar exam difficulty	Yes	See Table 1	Yes
Number of bar exam candidates	Yes	All states, 1985-2012	Yes
Candidates' mean MBE score	Only partially	12 states, selected years	No, *estimated* mean quality is used

Note: Candidates' mean quality (MBE mean score) is available for Alabama (1989-2001), Arizona (1991-2001), California (1992-2001), Colorado (1985-2001), Connecticut (1990-2001), Georgia (1991-2001), Maryland (1987-2001), Massachusetts (1985-2001), Missouri (1997-2001), Texas (1990-2001), Utah (1997-2001), and Virginia (1985-2001). See Appendix 1 for details on the estimation of candidates' mean quality.

Table 3. Differences between states with observable and unobservable exam difficulty.

	37 states with observable exam difficulty	14 states with no observable exam difficulty	
	Mean	Mean	Difference (p-value)
Number of candidates (log)	6.64	6.05	0.59 (0.06)
Pass rate	0.736	0.732	0.004 (0.85)
Number of candidates / state population (*1,000)	0.251	0.180	0.071 (0.02)
State population (/1,000,000)	6.03	3.53	2.50 (0.06)

Note: The table reports mean values for the 37 states with observable exam difficulty (Alabama, Alaska, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, D.C., Florida, Georgia, Hawaii, Illinois, Indiana, Kansas, Maine, Maryland, Massachusetts, Minnesota, Mississippi, Missouri, Montana, Nebraska, New Hampshire, New Jersey, New Mexico, New York, North Carolina, North Dakota, Ohio, Oklahoma, Pennsylvania, South Dakota, Texas, Utah, Virginia, and West Virginia) and the 14 states with no observable exam difficulty. The statistics are based on the 27 years between 1984 and 2011.

Table 4. Summary statistics (14 states that changed bar exam difficulty).

	Observations	Mean	Std. Dev.	Min	Max
Bar exam difficulty	370	134.5	4.25	125	144
Candidates' mean quality	370	141.6	3.33	133.1	150.5
Number of candidates (log)	370	7.14	1.28	5.11	9.65
Pass rate	370	0.728	0.104	0.374	0.921
Number of candidates / state population (*1,000)	370	0.253	0.146	0.118	0.804
State population (/1,000,000)	370	8.94	8.64	0.977	37.7

Note: The table reports summary statistics for the states with observable exam difficulty that changed exam difficulty at least once (Arizona, California, Florida, Georgia, Maine, Missouri, New Hampshire, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, and Utah). Table 1 describes changes in exam difficulty. Bar exam difficulty and candidates' mean quality are measured on a 0-200 scale.

Table 5. Summary statistics (23 states that did not change exam difficulty).

Variable	Obs	Mean	Std. Dev.	Min	Max	Difference (p-value)
Bar exam difficulty	423	133.8	3.69	128	145	0.771 (0.55)
Candidates' mean quality	423	140.9	3.87	129.2	151.7	0.744 (0.48)
Number of candidates (log)	423	6.33	1.21	3.53	8.31	0.813 (0.07)
Pass rate	423	0.732	0.099	0.410	0.943	-0.0038 (0.91)
Number of candidates / state population (*1,000)	423	0.264	0.204	0.053	1.825	-0.0105 (0.86)
State population (/1,000,000)	423	4.29	4.80	0.57	25.7	4.66 (0.08)

Note: The table reports summary statistics for the states with observable exam difficulty that did not change exam difficulty. Table 1 describes changes in exam difficulty. Bar exam difficulty and candidates' mean quality are measured on a 0-200 scale. The last column reports the difference in means between states that changed exam difficulty and states that did not (p-values in parentheses).

Table 6. The impact of potential labor supply on bar exam difficulty (OLS).

	(1) OLS Exam difficulty	(2) OLS Exam difficulty	(3) OLS Exam difficulty	(4) OLS Exam difficulty	(5) OLS Exam difficulty	(6) OLS Exam difficulty	(7) OLS Exam difficulty	(8) OLS Exam difficulty
Number of candidates (log), *t-1*	0.810*	0.835*	1.712**	1.642**	1.673**	1.857**	3.729**	3.586**
	(0.450)	(0.456)	(0.764)	(0.804)	(0.664)	(0.642)	(1.364)	(1.629)
Candidates' mean quality, *t-1*	0.651***	0.674***	0.315***	0.337***	0.990***	1.132***	0.490***	0.527***
	(0.137)	(0.155)	(0.0843)	(0.0849)	(0.128)	(0.175)	(0.0924)	(0.0766)
Year f.e.?		Yes		Yes		Yes		Yes
State f.e.?			Yes	Yes			Yes	Yes
Observations	793	793	793	793	370	370	370	370
Number of states	37	37	37	37	14	14	14	14

Note: The table reports OLS estimates of the impact of the number and quality of bar exam candidates on exam difficulty. In columns 5-8, states with no change in bar exam difficulty are excluded from the regression (see Table 1). Robust standard errors clustered by state are reported in parentheses.*** Significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Table 7. The impact of potential labor supply on bar exam difficulty (2SLS).

	(1) 2SLS Exam difficulty	(2) 2SLS Exam difficulty	(3) 2SLS Exam difficulty	(4) 2SLS Exam difficulty
Number of candidates (log), *t-1*	7.235*	11.34**	7.566***	13.50***
	(3.818)	(5.344)	(2.845)	(3.098)
Candidates' mean quality, *t-1*	0.0963	-0.0995	0.886***	0.435
	(0.0978)	(0.158)	(0.305)	(0.348)
Year f.e.?			Yes	Yes
State f.e.?	Yes	Yes	Yes	Yes
Observations	539	224	539	224
Number of states	36	14	36	14

Note: The table reports 2SLS estimates of the impact of the number and quality of bar exam candidates on exam difficulty. In columns 2 and 4, states with no change in bar exam difficulty are excluded from the regression. 2SLS estimates are obtained using as IVs the number of matriculated students and the mean LSAT score for students accepted into law schools in the same state in *t-4*. The number of observations is smaller than in Table 6 because of limited availability of data on the instrumental variables. In columns 1 and 3, the number of states is 36 because the instrumental variables are not available for Alaska. The first stage regression results are reported in Table 8. Robust standard errors clustered by state are reported in parentheses. *** Significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Table 8. First stage regression results.

	(1) OLS Number of candidates	(2) OLS Candidates' mean quality	(3) OLS Number of candidates	(4) OLS Candidates' mean quality	(5) OLS Number of candidates	(6) OLS Candidates' mean quality	(7) OLS Number of candidates	(8) OLS Candidates' mean quality
First year law school class size (/1000)	0.000242***	0.00154	0.000212***	0.00175*	0.000134**	0.00124	0.000149**	0.00131
	(3.93e-05)	(0.000929)	(2.25e-05)	(0.000931)	(5.79e-05)	(0.000960)	(6.62e-05)	(0.000982)
LSAT average score	-0.00603	0.468***	0.00561	0.676***	-0.0216*	0.246**	-0.00717	0.501**
	(0.00481)	(0.0762)	(0.00856)	(0.0977)	(0.0110)	(0.107)	(0.0215)	(0.191)
Year f.e.?					Yes	Yes	Yes	Yes
State f.e.?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	539	539	224	224	539	539	224	224
Partial R^2	0.06	0.21	0.10	0.40	0.06	0.04	0.06	0.14
F-test	18.89	20.49	54.02	27.09	6.40	3.06	3.26	3.46
(P-value)	(0.000)	(0.000)	(0.000)	(0.000)	(0.004)	(0.059)	(0.071)	(0.062)
Number of states	36	36	14	14	36	36	14	14

Note: Corresponding 2SLS results are reported in Table 6. In columns 3-4 and 7-8, states with no change in bar exam difficulty are excluded from the regression. The number of observations is smaller than in Table 5 because of limited availability of data on the instrumental variables. In columns 1-2 and 5-6, the number of states is 36 because the instrumental variables are not available for Alaska, see Appendix 1and 3 for more details on the instrumental variables and data availability. Robust standard errors clustered by state in parentheses. *** Significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Table A1. Additional results.

	(1) OLS Exam difficulty	(2) OLS Exam difficulty	(3) OLS Exam difficulty	(4) OLS Exam difficulty	(5) OLS Exam difficulty	(6) OLS Exam difficulty	(7) 2SLS Exam difficulty	(8) 2SLS Exam difficulty	(9) 2SLS Exam difficulty	(10) 2SLS Exam difficulty
Number of candidates (log), *t-1*			0.856*	0.887**	1.708**	1.624*	9.004*	14.21*	6.424*	14.90***
			(0.432)	(0.433)	(0.770)	(0.818)	(4.717)	(7.842)	(3.661)	(3.125)
Candidates' mean quality, *t-1*	0.738***	0.769***	0.740***	0.781***	0.316***	0.337***	-0.130	-0.760	0.184	-0.178
	(0.100)	(0.111)	(0.0976)	(0.110)	(0.0846)	(0.0849)	(0.243)	(0.623)	(0.130)	(0.351)
Admissions on motion, *t-1*			0.00155***	0.00172***	-6.12e-05	-0.000184				
			(0.000233)	(0.000259)	(0.000156)	(0.000186)				
Number of candidates / state population, *t-1*	7,338***	7,836***								
	(2,106)	(2,187)								
Year f.e.?		Yes		Yes		Yes				
State f.e.?					Yes	Yes	Yes	Yes	Yes	Yes
Observations	793	793	793	793	793	793	710	326	401	86
Number of states	37	37	37	37	37	37	36	14	31	9

Note: Columns 1 and 2 report the cross sectional correlation between potential supply and exam difficulty, where the number of candidates is divided by the population in the state. Figure 3 describes the simple correlation between these two variables (without controlling for candidates' quality). In columns 3-6, admissions on motion are the total number of admissions on motion in each state and year. In columns 7-8, the number of observations is larger than in Table 7, columns 1 and 2, because all observations from 1989 to 2012 are included in the sample (see footnote 27). Column 8 reports 2SLS estimates excluding states with no change in exam difficulty, as in Table 7, column 2. In columns 9-10, observations after the fourth year following each change in standards are excluded, this implies a decrease in the number of states in the sample. Column 10 reports 2SLS estimates excluding states with no change in exam difficulty, as in Table 7, column 2. Robust standard errors clustered by state are reported in parentheses. *** Significant at the 1% level, ** significant at the 5 percent level, * significant at the 10% level.

Table A2. Robustness results (different lags of the independent variables).

	(1) OLS Exam difficulty	(2) OLS Exam difficulty	(3) 2SLS Exam difficulty	(4) 2SLS Exam difficulty	(5) OLS Exam difficulty	(6) OLS Exam difficulty	(7) 2SLS Exam difficulty	(8) 2SLS Exam difficulty
Candidates' mean quality, [t-2, t-1]	0.412***	0.574***	0.0890	-0.119				
	(0.0899)	(0.0945)	(0.0895)	(0.191)				
Average number of candidates (log), [t-2, t-1]	2.485***	4.818***	7.126*	10.75**				
	(0.887)	(1.337)	(3.833)	(5.354)				
Candidates' mean quality, [t-3, t-1]					0.697***	0.573***	0.0938	-0.120
					(0.153)	(0.0993)	(0.0954)	(0.182)
Average number of candidates (log), [t-3, t-1]					0.845*	4.354***	7.284*	11.27**
					(0.447)	(1.311)	(3.835)	(5.499)
State f.e.?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	719	342	518	224	756	356	529	224
Number of states	37	14	36	14	37	14	36	14

Note: The table reports OLS and 2SLS estimates of the impact of the average number and average quality of bar exam candidates on exam difficulty. The average number of exam candidates and their quality is computed for [t-2, t-1] or [t-3, t-1]. The 2SLS estimates are obtained using as IVs the average number of matriculated students and the mean LSAT score for students accepted into law schools in [t-5, t-4] and [t-6, t-4] in columns 3-4 and 7-8 respectively. Robust standard errors clustered by state in parentheses. *** Significant at the 1% level, ** significant at the 5 percent level, * significant at the 10% level.

Table A3. Robustness results (OLS with state-specific and time-varying regressors).

	(1) OLS Exam difficulty	(2) OLS Exam difficulty	(3) OLS Exam difficulty	(4) OLS Exam difficulty	(5) OLS Exam difficulty	(6) OLS Exam difficulty	(7) OLS Exam difficulty	(8) OLS Exam difficulty	(9) OLS Exam difficulty	(10) OLS Exam difficulty
Number of candidates (log), *t-1*	1.543*	1.939**	3.628**	4.437***	1.536*	1.643*	3.406**	3.587**	0.858*	3.063**
	(0.831)	(0.892)	(1.587)	(1.396)	(0.778)	(0.818)	(1.523)	(1.630)	(0.424)	(1.110)
Candidates' mean quality, *t-1*	0.262***	0.302***	0.441***	0.516***	0.308***	0.337***	0.476***	0.527***	0.146*	0.265*
	(0.0717)	(0.0867)	(0.0750)	(0.0887)	(0.0783)	(0.0850)	(0.0822)	(0.0768)	(0.0803)	(0.123)
Real GDP per capita in state, *t-1*	0.0289**	-0.0315	0.0456*	-0.0686						
	(0.0141)	(0.0237)	(0.0217)	(0.0571)						
Manufacturing share of GDP in state, *t-1*					-9.090*	0.187	-10.49*	-0.0281		
					(4.675)	(5.946)	(5.288)	(9.997)		
Year f.e.?		Yes		Yes		Yes		Yes		
State f.e.?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-specific time trends?									Yes	Yes
Observations	747	747	337	337	793	793	370	370	793	370
Number of states	37	37	14	14	37	37	14	14	37	14

Note: The table reports OLS estimates of the impact of the average number and average quality of bar exam candidates on exam difficulty. Real GDP per capita in each state is an index equal to one in 2011; the share of manufacturing is computed using current GDP in each state (Source: Bureau of Economic Analysis). Robust standard errors clustered by state in parentheses. *** Significant at the 1% level, ** significant at the 5 percent level, * significant at the 10% level.

Table A4. Robustness results (2SLS with state-specific and time-varying regressors).

	(1) 2SLS Exam difficulty	(2) 2SLS Exam difficulty	(3) 2SLS Exam difficulty	(4) 2SLS Exam difficulty	(5) 2SLS Exam difficulty	(6) 2SLS Exam difficulty	(7) 2SLS Exam difficulty	(8) 2SLS Exam difficulty
Number of candidates (log), *t-1*	8.625**	7.542***	14.53**	13.71***	7.711**	7.575***	12.36**	14.16***
	(4.367)	(2.868)	(6.756)	(3.177)	(3.862)	(2.923)	(5.697)	(3.655)
Candidates' mean quality, *t-1*	0.122	0.886***	-0.179	0.451	0.0919	0.917**	-0.155	0.522
	(0.112)	(0.306)	(0.205)	(0.335)	(0.103)	(0.357)	(0.176)	(0.363)
Real GDP per capita in state, *t-1*	-0.0349	0.0163	-0.0528	-0.0217				
	(0.0301)	(0.0441)	(0.0438)	(0.0599)				
Manufacturing share of GDP in state, *t-1*					8.339**	-2.853	10.81	-7.451
					(3.771)	(10.21)	(7.664)	(17.35)
Year f.e.?		Yes		Yes		Yes		Yes
State f.e.?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	539	539	224	224	539	539	224	224
Number of states	36	36	14	14	36	36	14	14

Note: The table reports 2SLS estimates of the impact of the average number and average quality of bar exam candidates on exam difficulty. Real GDP per capita in each state is an index equal to one in 2011; the share of manufacturing is computed using current GDP in each state (Source: Bureau of Economic Analysis). Robust standard errors clustered by state in parentheses. *** Significant at the 1% level, ** significant at the 5 percent level, * significant at the 10% level.

DEPARTMENT OF ECONOMICS AND PUBLIC FINANCE “G. PRATO”
UNIVERSITY OF TORINO
Corso Unione Sovietica 218 bis - 10134 Torino (ITALY)
Phone: +39 011 6706128 - Fax: +39 011 6706062
Web page: http://eco83.econ.unito.it/prato/