

The Impact of Potential Labor Supply on Licensing Exam Difficulty

Mario Pagliero*

October 2011

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 show that a larger potential supply leads to more difficult licensing exams and lower pass rates. This implies that licensing partially shelters the market from supply shocks and limits the impact of policies targeted at increasing labor supply.

JEL: L4, L5, J44, K2.

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

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

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

fitness requirements. According to Kleiner (2000), over 800 occupations are licensed in at least one US state, including lawyers, accountants, auditors, teachers, nurses, engineers, psychologists, barbers and hairdressers. Occupational licensing directly affects 18 percent of U.S. workers, more than those affected by either minimum wage or unionization. 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 my 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 impact of the number of candidates on bar exam difficulty. Accounting for endogeneity increases the estimated impact of the number of candidates. Overall, the paper shows

¹I 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 US over the period 1981-2003. The grading procedures for the bar exam are described in Section 3.

that minimum entry requirements are relative standards, which respond to potential labor supply in the profession.

The magnitude of the impact of potential supply is considerable. Doubling the number of exam candidates implies an increase of about 8 percent in exam difficulty. This implies that the actual increase in labor supply is about half of the increase that would have happened without increases in standards. Thus, the licensing exam partially shelters 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 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

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

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 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 portion 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).

A related stream of literature has focused on the impact of licensing on wages and on the quality of professional services (Shepard 1978, HaasWilson 1986, Kleiner and Kudrle 2000, Kugler and Sauer 2005), labor mobility (Pashigian 1979), the origins of licensing (Law and Kim 2004), and the impact on minorities (Law and Marks 2009). Pagliero (2010) exploits changes in bar exam difficulty to estimate the impact of licensing requirements on entry-level salaries in the legal market. Harrington and Krynski (2002), and Harrington

³There 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).

(2007) study the funeral industry, Federman, Harrington and Krynski (2006) the market for manicurists, and Timmons and Thornton (2008) radiologic technologists. 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.⁴

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.⁵ It consists of 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.⁶

Essay and case questions are set by state boards and graded at the state level, according to criteria set by each board.⁷ In this case, a particular score does not necessarily correspond to a standard level of performance across states and years. However, most

⁴Pagliari (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.

⁵Exceptions are Delaware, Nevada and North Dakota, where the bar exam is held only once a year.

⁶A 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.

⁷Some 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 standards set locally.

states 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, Linn 1993).⁸

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.⁹ 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).¹⁰ 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 2010.¹¹ 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,

⁸An 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).

⁹They range between 0 and 200, like the MBE scores.

¹⁰The 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.

¹¹The 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.

consisting of MBE mean scores at the state level for each examination. Exam-specific 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.¹² 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. First, information on exam difficulty is not available for all the states (Table 1 reports information on 37 states). 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.

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.¹³ This problem is mitigated by using data on pass rates and exam difficulty to infer the mean MBE score for additional states and years. 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.

¹²Kleiner (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".

¹³Alabama, Arizona, California, Colorado, Connecticut, Georgia, Maryland, Massachusetts, Missouri, Texas, Utah, and Virginia.

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).¹⁴ 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 examinations, although the correlation is not statistically significant.¹⁵ 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 Table 2). 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.¹⁶ 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, and the main challenges to identifying the impact

¹⁴In 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.

¹⁵Similar results hold without dividing by the state population or dividing by the number of lawyers in the state.

¹⁶The 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).

of potential supply on exam difficulty.

4 Empirical results

I estimate regressions of the general form

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

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.¹⁷ λ_t and δ_i are state and 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 impact of potential supply on exam difficulty.¹⁸ Summary statistics are provided in Table 2.

Cross Sectional Evidence

Table 3, column 1-2 report OLS estimates of the pooled and cross sectional correlation between potential supply and difficulty of the bar exam. The impact of the number and the quality of candidates is positive and significantly different from zero. Adding year fixed effects in column 2 does not significantly affect the estimated coefficients. A ten percent increase in the number of candidates implies an increase of about 0.085 in exam difficulty (measured on the MBE scale). States with more candidates tend to have more difficult entry examinations.¹⁹ An increase of one in candidates' average quality implies an increase of about 0.7 in exam difficulty.

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.

¹⁷Estimates are accurate, and using the observed MBE scores, when available, does not affect the results (see Section 3.1 and Appendix 1 for details).

¹⁸Section 4.1 explores alternative specifications of (1).

¹⁹Similar results are obtained dividing the number of candidates by the population of the state (see Table A1, column 1).

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 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.²⁰

Within State Evidence

With these considerations in mind, I next focus on states that changed their entry standards in the sample period, so that state-specific unobserved characteristics can be accounted for by using state fixed effects. This implies a substantial reduction in sample size (from 37 to 13 states), but also a reduction in unobserved heterogeneity. To ensure that the reduction in sample size does not significantly affect the results, Table 3, columns 3-4 reproduce the results in columns 1-2 using the reduced sample. The impact of the number and quality of bar exam candidates is still positive and significantly different from zero, and the magnitude of the effects is even larger than before.

Table 3, column 5 reports the results controlling for state-specific fixed effects. The impact of the number and quality of bar exam candidates is still positive. The magnitude of the coefficients is quite different from that obtained using cross sectional variability. The impact of the number of candidates is 80 percent higher than that obtained in columns 3-4, while the impact 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

²⁰Similarly, 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.

supply and exam difficulty. Including both state and year fixed effects does not affect the estimated coefficients (Table A1, column 2).

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 worse 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, and an upward bias in the coefficient of quality. The 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).²¹

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).²²

Detailed information on class size for each law school in each year is available from the American Bar Association. Information on LSAT scores of admitted students provides a good measure of the candidates' average quality (the LSAT is a standardized examination used by law schools to screen applicants). The magnitude and statistical significance of these correlations can be estimated; hence the empirical relevance of the instruments can be tested.

Class size and quality of students admitted to law school are unlikely be affected by changes in exam difficulty three years later. Changes in standards are not typically

²¹A second reason for using instrumental variables is to account for possible measurement error in q_{it} .

²²See Oyer and Schaefer (2010) for evidence that law firms tend to hire largely from local law schools.

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 3, column 6 reports the 2SLS estimates of the impact of potential supply on exam difficulty. Table 4, columns 1-2 report the first stage estimates. As expected, column 1 shows that the number of 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, 11 and 38 percent respectively. The F-tests of the excluded instruments in the first stage are large, and the null of weak instruments can be rejected using the thresholds of Stock and Jogo (2002).²³

The 2SLS estimates in Table 3, column 6 show that the impact of the number of bar exam candidates on exam difficulty is again positive and significantly different from zero

²³The weak instrument problem is generally more relevant in overidentified than in just identified models.

at conventional levels. Its magnitude is larger than in column 5. Consider the impact 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.²⁴ In practice, however, the exam difficulty will grow in response to the increased potential supply by 11 (about 8 percent), which implies that only 36 percent of the new candidates will pass the exam.²⁵ Hence, about half of the effect on the number of entrants in the profession is offset by the increase in exam difficulty. The increase in the magnitude of the impact of the number of candidates is consistent with some remaining endogeneity in the within state regression (column 5). 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 estimate of the impact of candidates' quality is smaller than in column 5 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 far less precise than the corresponding OLS estimate (the standard error is double than in column 5) and the 95 percent confidence interval of the 2SLS estimate is $[-0.54, 0.21]$. Hence, on the one hand, the 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 interval still includes values which imply a significant impact of candidate quality on the number of entrants in the profession.²⁶

As with the OLS estimates, the results are not affected when both state and year

²⁴These 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 (see Appendix 2).

²⁵The 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 .

²⁶For example, 0.21 still implies that about 20 percent of the effect of increases in candidate quality on the number of entrants in the profession is offset by increases in exam difficulty (the effect is non-linear).

fixed effects are included (Table A1, column 3). Table A1, columns 7-8 report additional 2SLS results (discussed in detail in Appendix 2). Column 7 describes the results when the sample size is changed. Column 8 excludes observations after the fourth year following each change in difficulty to control for possible long-run responses of bar exam candidates to changes in difficulty.

4.1 Robustness Results

Although examinations are 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.²⁷ 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 4-6). The number of lawyers admitted on motion has 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 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.²⁸ If higher minimum MPRE score 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,

²⁷In 2004, there were fewer than 6,000 admissions on motion, but over 49,000 by examination (NCBEX). Only in the District of Columbia 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.

²⁸The 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).

all other variables being held constant. However, there are virtually no changes in MPRE minimum scores, so the results in Table 3, column 5 and 6 cannot be affected.²⁹

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.³⁰ 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.³¹ Moreover, the existing evidence suggests that these procedures directly affect only a very small number of candidates.

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 impact of the average potential supply in previous years. The results show a positive impact of potential supply on licensing exam difficulty.³²

5 Conclusions

Occupational licensing is one of the most important labor market institutions today, 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 impact of potential supply on the difficulty of the entry examination. The main finding is that supply considerations are important for explaining licensing requirements. In particular, increases in the number of potential entrants lead to more stringent entry

²⁹When 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.

³⁰Overall, 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.

³¹In 2005, 19 states have no published character and fitness standards in *The Comprehensive Guide to Bar Admission Requirements*.

³²For 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).

requirements, even after controlling for the average quality of candidates.

The existence of a relationship 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 between black 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 respond to market forces and, hence, are not uniquely based on some abstract notion of consumer protection, unrelated to market conditions.³³ 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

³³The 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).

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 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, it is unlikely that differences in asymmetric information across states are large enough to justify the large differences in exam difficulty observed in the data. It is not clear why asymmetric information should be more relevant in states with difficult exams (e.g., California) than 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

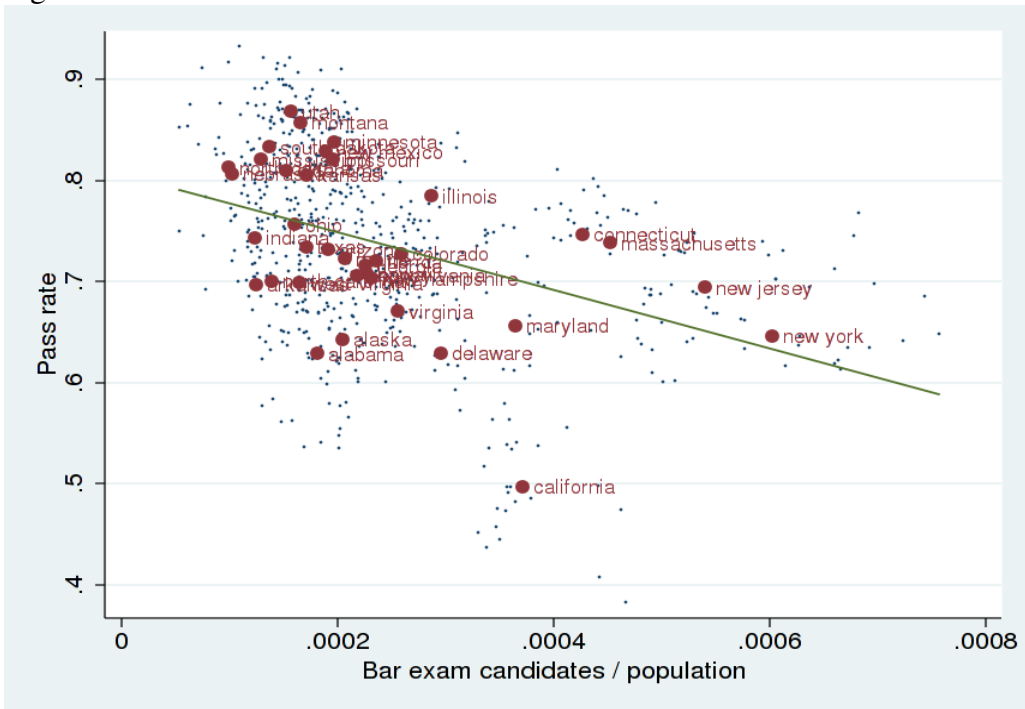
- Akerlof, G. (1970). The market for lemons: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics* 89, 488–500.

- Andrews, P. (2002). Self regulation by professions - the approach under EU and US competition rules. *European Competition Law Review* 23.
- 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, April). The Effect of State Funeral Regulations on Cremation Rates: Testing for Demand Inducement in Funeral Markets. *Journal of Law and Economics XLV*, 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. 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, 547–582.
- Kreuger, A. B. (2006, March 2). Do you need a license to earn a living? you might be surprised at the answer. *New York Times*, Economic Scene.
- 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 S. Kim (2004, May). Specialization and regulation: The rise of professionals and the emergence of occupational licensing regulation. Working Paper 10467, NBER.
- 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, 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/Employee Matching: An Industry Study of American Lawyers.
- Pagliari, M. (2010). Licensing Exam Difficulty and Entry Salaries in the US Market for Lawyers. *British Journal of Industrial Relations* 48, 4, 726–739.
- Pagliari, 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*, 1–25.
- 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.

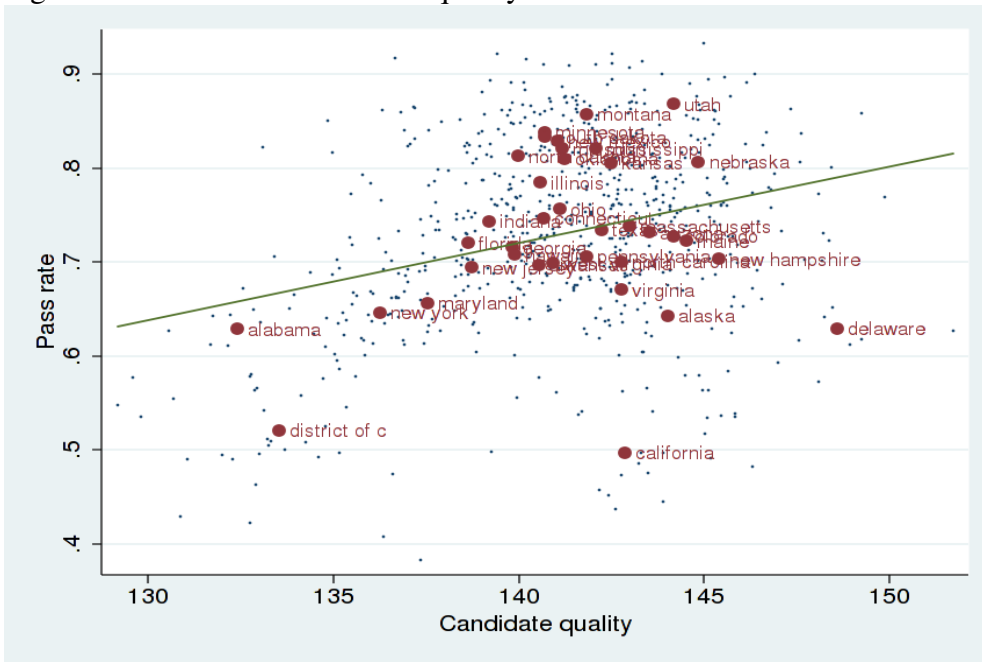
- Petersen, N. S., M. J. Kolen, and H. Hoover (1993). Scaling, norming and equating. In Linn (Ed.), *Educational Measurement*. American Council on Education / Macmillan.
- Shaked, A. and J. Sutton (1981). The self-regulating profession. *The Review of Economic Studies* 48, 217–234.
- Shapiro, C. (1986). Investment, moral hazard and occupational licensing. *The Review of Economic Studies* 53, 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 Rothenberg*, 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.

Figure 1 The Maurizi curve.



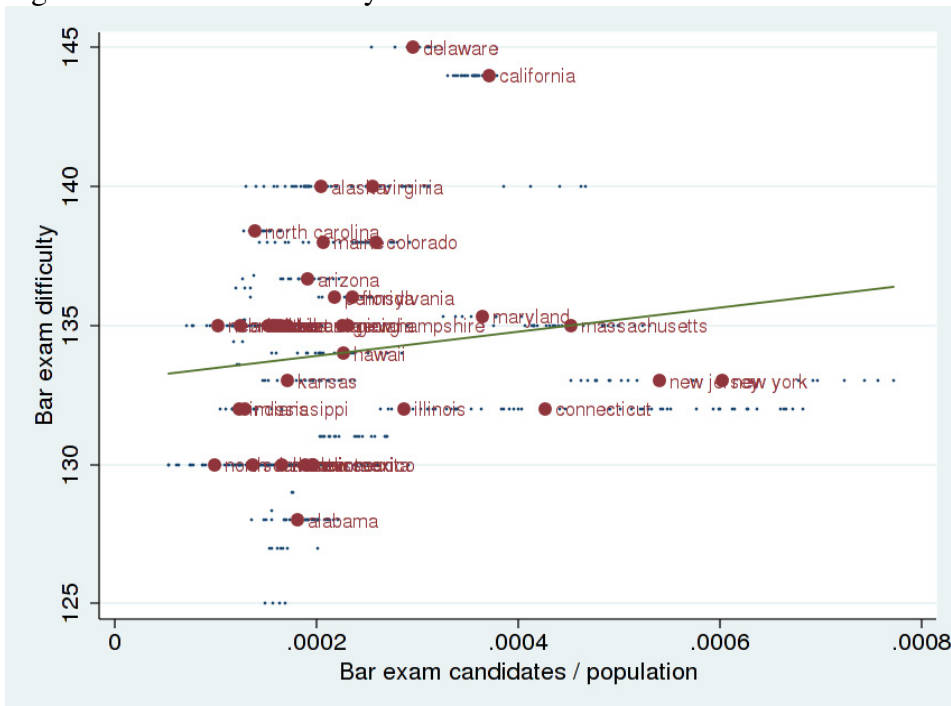
Note: DC is omitted from the graph as it is an outlier with large number of candidates (0.0086) and low pass rates (0.52). The linear correlation is statistically significant at 1 percent confidence level (and even more pronounced including DC). The blue (small) dots describe state-year observations. The red dots correspond to state-specific mean values.

Figure 2. Pass rate and candidate quality.



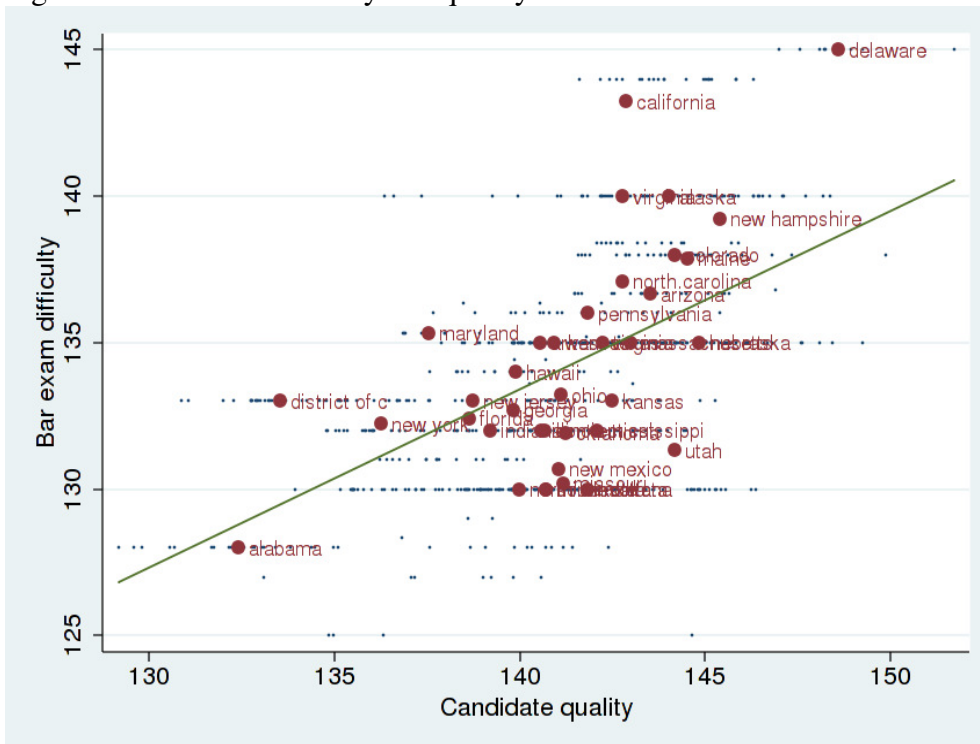
Note: the linear correlation is statistically significant at 5 percent confidence level. The blue (small) dots describe state-year observations. The red dots correspond to state-specific mean values.

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



Note: The linear correlation is not statistically significant. DC is omitted as it is an out-lier with large average number of candidates (0.0086). The correlation is unchanged 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 quality of candidates



Note: The linear correlation is statistically significant at 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 2010 (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	-	-	136.67
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 2010. Column 2 reports changes in difficulty, while Column 3 reports the corresponding date of each change. Column 4 reports difficulty in 2010. The information in Table 1 allows reconstruction of the time series of exam difficulty in each state.

Table 2. Summary Statistics.

Variable	Observations	Mean	Std. Dev.	Min	Max
Bar exam difficulty	719	134.07	4.00	125	145
Candidates' mean quality	719	141.10	3.62	129.17	151.71
Number of candidates (log)	719	6.72	1.31	3.53	9.62
Pass rate	719	0.73	0.10	0.37	0.93

Note: Bar exam difficulty and candidates mean quality (mean overall score) are measured on a 0-200 scale.

Table 3. The Impact of Potential Labor Supply on Bar Exam Difficulty.

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	OLS	OLS	2SLS
	Exam	Exam	Exam	Exam	Exam	Exam
	difficulty	difficulty	difficulty	difficulty	difficulty	difficulty
Number of candidates (log), $t-1$	0.822*	0.850*	1.702**	1.885**	3.477**	11.54*
	(0.450)	(0.456)	(0.664)	(0.643)	(1.507)	(6.479)
Candidates' mean quality, $t-1$	0.667***	0.692***	1.010***	1.149***	0.492***	-0.167
	(0.137)	(0.155)	(0.134)	(0.184)	(0.0963)	(0.191)
Year f.e.?		Yes		Yes		
State f.e.?					Yes	Yes
Observations	719	719	323	323	323	182

Note: The table reports OLS and 2SLS estimates of the impact of the number and quality of bar exam candidates on exam difficulty. Robust standard errors clustered by state are reported in parentheses. In columns 3-6, states with no change in bar exam difficulty are excluded from the regression (see Table 1). In column 6, the 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 columns 3-5 because of limited availability of data on the instrumental variables (see Appendix 1 and 3 for details on data availability). The first stage regression results are reported in Table 4. *** Significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Table 4. First Stage Regression Results.

	(1)	(2)
	Number of candidates	Candidates' mean quality
First year law school class size (/1000)	0.223***	1.45
	(0.034)	(1.01)
LSAT average score	0.00849	0.651***
	(0.00759)	(0.112)
State f.e.?	Yes	Yes
Observations	182	182
Partial R ²	0.11	0.38
F-test	25.50	19.59
(P-value)	(0.00)	(0.00)

Note: Corresponding 2SLS results are reported in Table 3, column 6. 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)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	2SLS	OLS	OLS	OLS	2SLS	2SLS
VARIABLES	Exam difficulty	Exam difficulty	Exam difficulty	Exam difficulty	Exam difficulty	Exam difficulty	Exam difficulty	Exam difficulty
Number of candidates (log), <i>t-1</i>		3.382*	15.99***	0.870*	0.905**	3.465**	14.59*	14.67***
		(1.715)	(5.197)	(0.431)	(0.431)	(1.514)	(8.863)	(4.322)
Candidates' mean quality, <i>t-1</i>	0.786***	0.516***	0.387	0.756***	0.799***	0.487***	-1.126	-0.535
	(0.112)	(0.071)	(0.318)	(0.097)	(0.111)	(0.093)	(0.971)	(0.519)
Number of candidates / state population, <i>t-1</i>	7,864***							
	(2,192)							
Admissions on motion, <i>t-1</i>				0.00154***	0.00172***	0.00141		
				(0.00022)	(0.00025)	(0.00185)		
Year f.e.?	Yes	Yes	Yes		Yes			
State f.e.?		Yes	Yes			Yes	Yes	Yes
Observations	719	323	182	719	719	323	279	154

Note: Column one reports the cross sectional correlation between potential supply and exam difficulty, as in Table 3, column 2, but the number of candidates is divided by the population in the state. Column 3 reports 2SLS results with state and year fixed effects. In column 3, Newey-West standard errors are reported in parentheses (Bartlett kernel function, with bandwidth 3). Clustered standard errors are similar but unreliable due to the drop in degrees of freedom (relative to Table 3, column 6) caused by the inclusion of state fixed effects. This implies that standard errors in column 3 are not directly comparable with those reported in Table 3, column 6. In columns 1-2 and 4-7, robust standard errors clustered by state are reported in parentheses. Column 7 reports 2SLS estimates, as in Table 3, column 6: The number of observations is now larger (279) than in Table 3, column 6 because observations from 1989 to 2010 are used (see Appendix 2 for details). Column 8 reports the results using the same sample as in Column 7, but excluding the observations after the fourth year following each change in standards (see Appendix 2 for details). *** Significant at the 1% level, ** significant at the 5 percent level, * significant at the 10% level.

Table A2. Robustness results.

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	2SLS	OLS	OLS	2SLS
	Exam	Exam	Exam	Exam	Exam	Exam
	difficulty	difficulty	difficulty	difficulty	difficulty	difficulty
Candidates' mean quality, [t-2, t-1]	0.732*** (0.163)	0.584*** (0.104)	-0.204 (0.232)			
Average number of candidates (log), [t-2, t-1]	0.878* (0.451)	4.454** (1.559)	10.10* (5.617)			
Candidates' mean quality, [t-3, t-1]				0.711*** (0.153)	0.574*** (0.106)	-0.190 (0.210)
Average Number of candidates (log), [t-3, t-1]				0.861* (0.448)	4.007** (1.466)	10.87* (6.109)
State f.e.?		Yes	Yes		Yes	Yes
Observations	646	298	182	683	311	182

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 column 3 and 6 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.

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.³⁴

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.³⁵

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 some assumptions on the score distribution.

³⁴ 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.

³⁵ The results are consistent with what I would obtain using the original data, but they are much more precise.

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

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

where $F(\cdot)$ 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

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

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 .³⁶

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),

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

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

More details on the estimation are available in Appendix 3.

³⁶The results are not significantly different when I assume that scores have a beta distribution, which can accommodate some skewness in the score distribution.

Appendix 2. Instrumental Variables

The two instrumental variables are the LSAT mean score and the number of matriculated students in law schools for each state and year. Information on LSAT scores of admitted students for each law school is available from the ABA Official Guide to Law Schools. As the LSAT is a standardized examination, results can be compared across states and years. Detailed information on class size is also available from the same source. Both LSAT score and class sizes are aggregated at the state-year level (average LSAT scores are appropriately weighted).

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 usable observations in Table 3, column 6 is significantly smaller than for the 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, column 7). This suggests that the shorter time series used for the 2SLS results in Table 3 are not driving the results.

Table A1, column 8 excludes the observations after the fourth year following each change in standards. The results are not affected. This 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). Hence, results cannot be driven by a long run endogenous response of bar exam candidates to changes in difficulty. When more than one change in exam difficulty occurs in a given state, all the observations after the fourth year following the first change are dropped.