Teacher Licensing in U.S. Public Schools: The Case for Simplicity and Flexibility

Michael Podgursky
Department of Economics
University of Missouri—Columbia

Discussions of teacher quality in U.S. public schools have tended to focus on teacher licensing. Reformers from diverse perspectives have proposed “raising the bar” for licensing by various means (e.g., harder licensing exams, more coursework, graduation from the National Council for Accreditation of Teacher Education teacher-training programs, elimination of emergency or provisional licensing). However, all of these proposals assume that teacher licensing plays an important role in determining teacher quality and performance. In this article I argue the contrary. Raising the bar for teacher licensing in ways that have been proposed is unlikely to have any significant short- or long-term effects on student achievement. Moreover, by shrinking the applicant pool for vacancies, these restrictive proposals may have the perverse effect of lowering average teacher quality, particularly for high-poverty or rural districts that already face thin applicant pools. A preferred approach is to swap regulation of inputs for accountability for outputs, that is, a more

An earlier version of this article was presented at the Teacher Preparation and Quality: New Directions in Policy and Research Conference, American Enterprise Institute, Washington, DC, October 18–20, 2003. I thank Youn Soel and Erin Allen for research assistance, Michael Wolkoff and conference participants for many thoughtful comments, and the Smith Richardson Foundation for financial support. The usual disclaimers apply.

Requests for reprints should be sent to Michael Podgursky, Department of Economics, University of Missouri—Columbia, 118 Professional Building, Columbia, MO 65211. E-mail: Podgurskym@missouri.edu
flexible licensing regime that relaxes entry barriers combined with greater accountability for student achievement gains. Such a proposal is based on a simple economic principle: Welfare is more likely to be improved if state regulators focus on what they can measure (student achievement), not what they cannot (teacher quality).

Introduction: Six Propositions About Teacher Labor Markets

Teacher licensing figures prominently in popular discussions of teacher quality. On the one hand, teachers unions, the American Association of Colleges of Teacher Education, and the National Commission on Teaching and America’s Future (NCTAF) have pressed to make teaching more like medicine with national teacher testing standards, mandatory accreditation of teacher training institutions, “independent” state teacher licensing boards, and similar entry restrictions (Ballou & Podgursky, 2000; NCTAF, 1996, 1997). Critics from outside of the mainstream education community have also targeted the teacher licensing process and called for more rigorous exams and stronger academic training (Ravitch, 1998; Walsh, 2004). Teacher licensing also figures prominently in the teacher quality standards of the No Child Left Behind Act. Among other things, states are required to phase out emergency and other types of “substandard” teacher licensing. Finally, shortfalls of fully licensed teachers figure prominently in school finance “adequacy” litigation where plaintiffs make the case that school spending is inadequate based on the higher incidence of inappropriately licensed teachers in low-income, high-minority plaintiff school districts (e.g., Darling-Hammond, 2003). Such arguments clearly played a role in the recent New York State court decision overturning the state school finance system (Campaign for Fiscal Equity v. New York State).

A common theme in all of these examples is a belief that teacher licensing plays an important role in determining teacher quality and performance. In this article, I argue that “raising the bar” in teacher licensing, based on the criteria advocated by reformers, is unlikely to have any significant short- or long-term effects on student achievement. Teacher quality in schools has much more to do with personnel policies such as incentive pay, performance reviews, dismissals, and effective professional development. Moreover, because various proposals to raise the bar shrink the applicant pool for available vacancies, they may have the perverse effect of lowering average teacher quality, particularly for high-poverty or rural districts that
already face thin applicant pools. A preferred approach is a more flexible licensing regime that would relax entry barriers and permit school districts greater discretion in hiring decisions in exchange for greater accountability for student performance. To lay the groundwork for this argument, I begin with several propositions concerning research on teachers and the economics of occupational licensing.

*Research Linking Teacher Training or Licensing to Student Achievement Is Inconclusive and Provides Little Support for Aggressive Regulation of the Labor Market*

If policymakers choose to raise the bar or more rigorously enforce teacher licensing, they should have solid evidence that the criteria used to exclude teacher candidates from the market has a demonstrable relation to student achievement. Otherwise, such policies will simply shrink the size of teacher applicant pools without raising the average quality. How strong is the research base for formulating policy concerning teacher training and licensing? Recent surveys of the scientific research base find it is very thin.

Such an assertion seems to fly in the face of claims by various education groups about the “knowledge base” for teaching and certainly seems contradicted by the hundreds of studies published annually in education research journals, many of which are devoted to teacher education. If, however, there is widespread consensus in the social science research community that scientific evaluation of social policy programs (including education) requires (a) randomized experimental study design or (b) nonexperimental longitudinal data on participants. Unfortunately, little research on teacher testing, training, or licensing meets either standard, and the research that does is tentative and inconclusive.

1 For example, a 1996 report of the NCTAF, a self-appointed commission including the president of the National Council for Accreditation of Teacher Education and the National Board for Professional Teaching Standards, stated, “Although hundreds of studies have shown that fully prepared teachers are more effective than those who are unqualified, the practice of hiring untrained teachers continues” (p. 15). “Teachers who know how to do these things [pedagogy] make a substantial difference in what children learn. Furthermore, a large body of evidence shows that the preparation teachers receive influences their ability to teach in these ways” (p. 27).

2 A third model, most frequently used by economists, is the “natural experiment” model (Heckman, Lalonde, & Smith, 1999). This has not been widely employed in the teacher literature. An exception is Jacob and Lefgren (2002), who use longitudinal student-level achievement data for Chicago public school students and exploit a quirk in the administrative regulation to create a “quasiexperiment” to examine the effect of teacher training on student achievement. However, even with natural experiments, longitudinal data are highly desirable.
Randomized experimental design is the “gold standard” for social policy research. With respect to teacher quality, this would involve estimating the effect of teachers with different credentials or training on student achievement through random assignment of students to classrooms of variously credentialed but otherwise comparable (e.g., experience) teachers within a school. I am aware of only one study of teacher training or credentials that meets this standard, although the Institute for Education Sciences of the U.S. Department of Education is actively promoting such studies (Mosteller & Boruch, 2002; U.S. Department of Education, n.d.), and more are under way. Thanks to these efforts, 5 years from now it is likely that we will have experimental evidence on teacher licensing and training.

If randomization is not feasible, and often it is not, then one must rely on nonexperimental data to evaluate education policy. If we are to measure the contribution of a classroom teacher to student achievement, it is necessary to control for prior achievement of the student before he or she enters the classroom. Ideally, researchers would pretest the students in the fall and test them again in the spring. The difference in these scores, averaged over the classroom, would be a measure of a teacher’s “value added.” If students are not pretested in the fall, then it is also possible to use test scores the previous spring, or for more than 1 previous year (longitudinal achievement data). Large longitudinal data files have formed the basis for the most sophisticated current research on teachers and teacher effects on student achievement (Aaronson, Barrow, & Sander, 2003; Rivkin, Hanushek, & Kain, 1998; Sanders & Horn, 1994).

Studies that do not have a rigorous study design (i.e., with randomization or controls for prior student achievement) are likely to produce seriously biased estimates of the effect of teacher certification or other teacher characteristics on student achievement. The reason is that they do not adequately control for the socioeconomic status (SES) background of students in classrooms, and student SES is correlated with teacher credentials and strongly correlated with student achievement. In the language of econometrics, cross-section studies of the effect of teacher credentials on student achievement suffer from “omitted variable bias.”

A recent study of Teach for America (TFA) conducted by Mathematica Policy Research (Decker, Mayer, & Glazerman, 2004) used random assignment methods. They found that students taught by TFA teachers had significantly larger achievement gains in math. Estimates of achievement gains in reading were also larger, but the difference between TFA and traditionally trained teachers was not statistically significant.

A recent study by Hoxby (2001) highlighted the importance of these socioeconomic variables and their potential for producing bias in studies of teachers and student achievement. Hoxby analyzed the effect of family, neighborhood, and school input variables on student
For example, consider studies of the effect of teacher licensing on student achievement. Given the complexities of teacher licensing systems, virtually every school district in the United States has some teachers out of compliance; however, substandard certification tends to be relatively more common in schools with low SES students. Because SES has a very powerful effect on student achievement levels and gains, unless the researcher has very good controls for prior achievement and SES in a study of certification and student achievement, the resulting study is likely to yield an upward-biased estimate of the effect of certification.5

The number of studies of teacher certification that meet the minimum methodological standards outlined earlier is very small. A recent survey of the literature by Wayne and Youngs in the spring 2003 Review of Education Research found only two studies of teacher certification that were peer reviewed, used longitudinal student-level achievement data, and controlled for student SES. The results of these studies (both by Goldhaber & Brewer, 1997a, 1997b, and both using the National Longitudinal Educational Survey of 1988) had mixed results. They did find a small positive effect of math teacher certification on math achievement, but no statistically significant effect of science teacher certification on science achievement. Recent surveys of the literature by Hanushek and Rivkin (2004) focusing on “high-quality” studies that meet the standards described earlier find little evidence linking teacher credentials to student achievement. For example, of nine estimates of the effect of teacher test scores on student achievement, six found no statistically significant

achievement and educational attainment using two large nationally representative longitudinal studies of students (the National Educational Longitudinal Survey [NELS], NELS88, and the National Longitudinal Survey of Youth, which began in 1979). Hoxby compared the percentage of the variation in student achievement on various field tests (math, reading) explained by school, family, and community factors. For every test, the percentage of the variation explained by the family variables far exceeded the school input variables. The family variables explained from 34 to 105 times as much variation in student achievement test scores as the school input variables. She also examined years of schooling completed at age 33. Family variables explained 19 times as much variation in student educational attainment as did school inputs.

5Moreover, this is not a problem that is “fixed” by meta-analyzing large numbers of flawed cross-section studies because all of these studies are biased in the same direction. Meta-analyzing 200 such studies simply produces a more accurate estimate of a biased coefficient. A target shooting analogy can illustrate this point. If the scope on a rifle is off or out of adjustment (biased), then the rifle shots will cluster around a point that is away from the target bull’s-eye. Firing more shots will simply do a better job of identifying the point around which the sight is targeted but will not help determine where the bull’s-eye is. That requires that the bias or error in the rifle scope be fixed.
effect. Of the three finding a significant effect, two were positive and one was negative.\(^6\)

In short, the research foundation for raising the bar with teacher tests or raising standards for schools of education is weak. The evidence linking any type of teacher training, licensing, or testing to student achievement is mixed at best. Even estimated effects of general academic skills of teachers such as SAT scores, although usually statistically significant, are generally modest in effect.

**Teacher Effects on Student Achievement Are Quantitatively Important but Idiosyncratic**

Does this mean teachers do not matter? On the contrary, although the effect of measured teacher characteristics is small, one consistent finding is that there seems to be considerable variation in teacher effectiveness between classrooms. Therefore, if one compares the effect on student learning of the top and bottom 20% of teachers ranked by performance, the effect is often quite substantial. However, these teacher effects are largely unrelated to traditional measures of teacher quality such as licensing exam test scores, certification credentials, experience, or graduate degrees, a result highlighted in a recent survey by Goldhaber (2002). Hanushek and Rivkin (2004), summarizing their own and other research, came to the same conclusion.

A recent study by Aaronson et al. (2003) of Chicago public teachers illustrates this point well. Like other such studies, this work was based on a very large longitudinal file of linked student achievement scores. What makes this study unique is that the authors also had extensive administrative data on teacher characteristics that are unavailable in other studies, including education, experience, types of teaching licenses, and selectivity of

\(^6\)A recent survey of teacher quality research by the Education Commission of the States (Allen, 2003) sets a lower standard for inclusion of studies. Allen considered cross-section as well as descriptive studies. Nonetheless, he found at best tepid research support for aggressive regulation of the teacher labor market. On the question of whether pedagogical training contributes to teacher effectiveness, he found only “limited” support in the research, and added “It is not clear from the research reviewed for this report, however, whether such knowledge and skills are best acquired through coursework, field experience (especially student teaching) or on the job” (p. 29). On the question of whether more stringent screening for teacher training program entrants pays off in terms of student achievement, he found the literature “inconclusive.” A new study by Betts, Zau, and Rice (2003) examined student achievement gains in the San Diego school district. They found mixed results for teacher credentials. In some cases, students of emergency certified teachers have higher gains than those of experienced fully credentialed teachers. At the upper grade levels, full certification in math has a significant positive effect in high schools but a negative effect in middle schools.
the teacher’s undergraduate college. They found that over 90% of teacher effects are not explained by any measured teacher characteristics.

In sum, the growing “teacher effects” literature suggests that teacher quality, as measured by student achievement gains, is highly idiosyncratic. This does not mean that teacher quality is random or unknowable. It simply means that traditional measures of teacher quality such as experience, master’s degree, and education coursework explain virtually none of the variation in teacher effectiveness.7

In the Absence of Strong Ex Ante Indicators of Teaching Quality, Raising the Bar in Teacher Licensing Is Likely to Lower Teacher Quality

A skeptic might argue that, although the evidence for teacher licensing, testing, or a particular program of pedagogical training is weak, why not raise the entry bar anyway on the chance that it might work? Many public policies are enacted on faith and good intentions rather than rigorous scientific research. Why is teacher licensing any different? What harm can come from raising the bar for teachers?

If these reforms were costless, then one might make the case for their implementation on the chance that some benefits would accrue. However, they are not costless, and there is a very real possibility that schools will find themselves worse off and student achievement will fall if such programs are implemented.

First, there are the direct resource costs. To the extent that we raise requirements for education coursework, we incur direct educational costs as candidates take classes, pursue professional development, forego other employment, and so forth. Annual costs per student in higher education currently are roughly $27,000 dollars per year, although students on average only pay part of this cost. Tests are less costly, but the fixed costs of updating and validating new teacher tests is considerable. More important are the time costs for teaching candidates spent in pedagogy courses or preparing for and taking exams. If we assume teaching candidates take 1½ years of teaching courses (including student teaching), this is a very costly investment. Even at the minimum wage, this amounts to over $15,000. One perverse result of requiring seat time in pedagogy courses for labor market entry is that candidates with greater academic skills, who presumably

7It may be that other potential measures, involving direct observation of classroom practice, or psychological assessments of teacher attitudes toward students and teaching may do a better job of explaining the classroom effectiveness. Indeed, many school districts use the latter types of assessments in their hiring decisions.
have a higher alternative wage, face a higher cost in securing a teaching license (Ballou & Podgursky, 1997).

Only roughly 60% of all teacher candidates graduate from NCATE-accredited teacher training programs. Closing all teacher training programs that do not secure NCATE accreditation, as proposed by the National Education Association and the NCTAF, would almost surely restrict the flow of newly trained teaching candidates. So, too, would raising the cutoff on teacher licensing exams. If the cutoff is raised from the 20th to the 30th percentile on elementary education exams, then 10% of potential applicants are excluded from the applicant pool.

A common feature of all of these policies is that they would reduce the applicant pool to public schools. Other things being equal, this will tend to lower the average quality of teachers who are hired. Why is this the case? School administrators know many things about teacher candidates that state regulators do not. They conduct job interviews, evaluate student teaching, read letters of recommendation and transcripts, and observe demonstration classes. In fact, school administrators are in a much better position to assess teacher quality than are state regulators, and there is some evidence that their assessments can identify teachers who produce larger student achievement gains. By preventing school administrators from considering any unlicensed applicants, school districts are forced to hire the worst certified candidate even if a superior noncertified candidate is available. Raising the bar shrinks the size of the applicant pool that school administrations may consider. The new pool is better in terms of whatever the regulators specify (e.g., more NCATE graduates, higher Praxis II scores), but now the administrators have fewer candidates in the applicant pool. Therefore, they have less ability to select among candidates on the basis of the factors that they observe, but state regulators do not.

This cost of mandatory certification is illustrated in Figure 1. Here I have presented hypothetical data on the distribution of teacher quality among certified and uncertified applicants. Although these data are hypothetical, I believe that they represent the general picture that is emerging in the teacher effects literature—namely, that the individual variation in classroom performance of teachers is large relative to any measurable teacher characteristic such as teacher certification. As I indicated earlier, the evidence concerning teacher certification is mixed at best. However, for the sake of argument, I have assumed a positive effect: The average certified

8Studies using student longitudinal data by Armor et al. (1976) and Murnane (1975) find large effects of principal evaluations on student achievement gains. More recently, Sanders and Horn (1994) reported, “There is a very strong correlation between teacher effects as determined by the data and subjective evaluations by supervisors” (p. 300).
teacher is better than 60% of noncertified teachers. Based on the review of the research literature discussed earlier, I see this as an upper-end estimate for certification. However, the conclusions that follow do not hinge in any significant way on this assumption. If we bumped the assumption upward and had certified teachers better than 80% of potential candidates, none of our basic conclusions would change. The key point is that the research suggests that there is a large dispersion of quality within the certified and uncertified pools.

Suppose Figure 1 represents the population from which a school district recruits teaching candidates. Further suppose that the school district has a single vacancy and is free to hire the best candidate, certified or not. Imagine that a single candidate applies at random from the certified pool and one applies from the uncertified pool. What is the probability that the certified candidate is the superior teacher? In Table 1 we see that 57% of the time the certified candidate is better. With 2 job applicants, the average quality of the best teacher (certified or not) is at the 67th percentile of the certified distribution. Now suppose the school has 5 random applicants from the certified population but no uncertified applicants. Note that the quality of the best applicant jumps sharply from the 67th to the 88th percentile. This illustrates an important point. If teachers are screened well (a point taken up later), a larger applicant pool means better quality hires. The reason is that you are hiring the best applicant out of a pool, not the average applicant; therefore, it is better to have more rather than few appli-

![Figure 1. Overlapping ability when the average certified applicant is better than 60% of noncertified applicants.](image)
cants. This common sense point is borne out in many other contexts. Other things equal, the average quality of graduate students will be higher in a program with 200 applicants as compared to one with 20, even if the mean quality of the two applicant pools is the same.

Can this district similarly benefit from access to a pool of lower average quality uncertified applicants? The answer is yes (see bottom of Table 1). Here we suppose that 5 uncertified applicants are allowed to apply for the job along with 5 certified applicants. Although the uncertified applicants are of lower average quality than the certified candidates, 39% of the time the uncertified applicant will be the best of the 10 applicants. Why? The last column shows that by expanding the applicant pool from 5 to 10 candidates, the mean quality of the best teacher has increased as well. Raising the bar and restricting the applicant pool moves us in the opposite direction. If we go from row 3 with 10 applicants to row 2 with 5 applicants, average teacher quality falls despite the fact that the average certified teacher is better than the average noncertified teacher. The fall in mean teacher quality as we go from row 3 to row 2 illustrates the hidden cost of a licensing entry barrier: Shrinking the applicant pool gives schools fewer choices, gives them less freedom to pick out talent, and reduces the mean quality of the resulting hires.9

To summarize, raising the bar on teacher licensing shrinks the applicant pool available to schools. Other things being equal, this will tend to lower

9A skeptic might argue that I have “stacked the deck” in this simple simulation by assuming that in hiring the school district screens perfectly and always hires the best candidate. However, in a more elaborate simulation, Ballou (1999) assumed that school administrators have imperfect, but independent, information about the quality of job applicants (i.e., over and above certification or a test score). Such information might include direct observation of teaching, recommendations, or prior teaching experience. In that case, we find a similar result to the one we observe here. Ballou also showed that the cost of the reduced supply tends to be larger for low SES districts that tend to draw relatively more applicants around the cut scores.
the quality of the teaching workforce. This negative effect or cost may be offset if licensing raises the mean quality of the (smaller) applicant pool. However, the research to date provides little solid evidence for such a positive effect. Therefore, we are faced with the likelihood that mean teacher quality will actually decline as a result of such policies.10

Teaching Is Not Medicine

Whether or not a research base currently exists to support aggressive licensing of teacher labor markets, proponents often argue that teacher “professionalization” is a desirable end in itself. They appeal to a vision of professional self-regulation in education akin to that in medicine. In panel discussions on teacher licensing, I am routinely confronted with the question, “Would you send your children to an unlicensed doctor?” Although rarely stated explicitly, I believe the argument runs as follows: Although it may be true that there exists no rigorous evidence for the reforms we have proposed (e.g., tougher teacher testing, accreditation, and more vigorous review of teacher training institutions), they are broadly similar to what is found in medicine. Therefore, if we implement such reforms, teacher quality and the quality of education will improve in the manner seen in medicine.

Argument by analogy is valid only if the analogy is valid. Why teaching is not medicine deserves an entire article of its own; however, I briefly explain two important reasons why this analogy is inappropriate.

There is a deep body of scientific research in medicine, and commitment to scientific research methods pervades medical schools, the professional specialty associations, and the community of medical practitioners generally. The economic case for medical licensing rests on an information asymmetry between what these highly trained medical practitioners know and what consumers know concerning the quality of the services they are buying. Because of the complexity of the knowledge base in medicine and the high cost of mistakes (malpractice), it is relatively easy to argue that some sort of government licensing is required to screen out incompetent practitioners and protect consumers.

Does this model apply to K–12 education? If we replicate the professional self-regulation found in medicine, can we expect qualitatively similar outcomes in education? To be sure, there is scientifically based research on student learning. However, for the most part, this research is being produced not in schools of education but by educational and cognitive psychologists in psychology departments.

Ballou (1999) showed that this negative net effect is more likely to occur in high-poverty districts where many teachers hired are near the cut scores on licensing exams.
Even at leading research universities, the majority of education school faculty do not produce research based on rigorous scientific methodology—certainly nothing akin to what one finds in a medical school. Many education faculty approach research with methods more akin to those found in the humanities than those in medicine or the sciences. Controlled experiments and randomized studies are rare. Use of large-scale longitudinal data on students is not widespread. However, what education school faculty at leading research universities do or do not do is largely irrelevant because they train relatively few of the nation’s classroom teachers. The primary supplier of classroom teachers (as opposed to doctors) are state colleges—most of which were formerly teacher’s colleges. At such institutions, much of the teaching is conducted by adjunct faculty not actively engaged in scientific research. To the extent that regular faculty at such institutions do research at all, it cannot be described as scientifically rigorous and is far removed from the frontiers of scientific research on human learning. The same can be said of other areas of education policy research.

However, even if upper and lower tier schools of education were producing scientifically based research, the teachers and their professional associations are in no position to vet this research and incorporate it into their teaching or their standards. Professional teacher associations such as the National Council of Teachers of English or the National Council of Social Studies do not base their standards on scientific research. Indeed, most members of these learned societies are practicing teachers and not trained to evaluate scientific research. I venture that most practitioners and education school professors in these fields would not even view the scientific method (i.e., experiments) as the most useful method of inquiry in their field.

The deep technical and scientific knowledge base in medicine produces well-defined and widely shared agreement on appropriate clinical practice. For the most part, this is absent in education. Although the judgment of English, mathematics, and elementary school teachers as to the best ways of teaching a subject certainly deserve respect and deference, there is little evidence to suggest that parents cannot make informed choices among practitioners who approach their craft differently.\textsuperscript{11} This leads us to the next proposition.

\textsuperscript{11} A recent paper by Angrist and Guryan (2003) found that states with teacher testing have higher teacher pay (indicating a restriction in supply) but no higher teacher quality as measured by various academic quality indicators (e.g., selective college graduates). Another subtle difference between teaching and medicine (as well as other professions) deserves mention. In medicine, the primary desire of a patient is simply to be made well. When we go to a doctor...
Unregulated Markets in Education and Training Work Well

The case for the medical analogy would be strengthened if there were pervasive evidence of “market failure” in unregulated markets for education and training. Although I am not aware of widespread unlicensed practice of surgery, unlicensed training and schooling is pervasive in our economy. The latter markets seem to work quite well with little or no government regulation. In fact, a review of the functioning of these labor markets suggests that they operate considerably better than the highly regulated markets in public K–12 education.

Researchers have estimated that American business spends between $18 and $43 billion (1995 dollars) annually on formal training programs for their workers and an unknown but substantial amount on informal training (Ehrenberg & Smith, 1996, p. 302). Virtually all of this training is delivered by instructors who are not licensed by the state and who have not received specialized pedagogical instruction. Historically, one of the most important sources of high-quality vocational training in our economy has been the U.S. military. The various services have taken millions of high school dropouts and graduates and provided them with high-quality training in technical specialty fields. Along the way, in the process of turning millions of young men and women with limited elementary and secondary education into trained aircraft mechanics, radio operators, supply clerks, and so forth, the armed services have taught these young recruits basic literacy and numeracy skills as well. Nearly all of this was accomplished by unlicensed instructors.

with a ruptured appendix, a dentist with a toothache, or a lawyer for legal representation, we want a “sage on the stage,” not a “guide at the side.” That is, we want their professional expertise put to work solving our problem. Usually, the process is a secondary concern to the end, and we usually defer to the judgment of the expert professional on the best course of “treatment.” Of course, if there are several ways to achieve the same end, the consumer will need to make a choice. However, more often than not, the treatment protocols are standard, and the consumer follows the advice of the doctor to achieve the desired end (a cure). However, in education, for many parents, the process is as important as the end result. Indeed, the two can be hard to separate. When parents choose a Montessori or a Waldorf school for their children, they clearly expect their children to learn basic literacy and numeracy skills, but they are also expressing a preference over a mode of inquiry and learning as well. Similarly, when parents object to the use of calculators by young children as in the initial National Council of Teachers of Mathematics standards or to whole language reading instruction, they are expressing a preference for a type of instruction as well as an outcome. In fact, the experience in the private K–12 education marketplace suggest that parents are perfectly capable of making informed choices among vendors who offer a wide range of instructional strategies (e.g., from constructivist, to traditional, to military schools) and can select a school that meets their preferences. We see little evidence of market failure or calls for government regulations coming from private school consumers.
Approximately 6 million students are enrolled in 2-year community colleges. Much of the coursework offered in these community colleges is remedial and covers material that students should have learned in elementary and secondary schools. States do not require the faculty in community colleges to be licensed, and evidence suggests that most are not certified teachers. Nonetheless, if we judge success by enrollment growth, successful transition to 4-year baccalaureate institutions, or higher earnings, these community colleges are successfully delivering K–12 educational services.

Many students receive K–12 educational services from the thousands of private tutoring firms. These range from large multinational educational firms like Sylvan Learning to small independent proprietary firms. Many of these firms specialize in providing remedial help for students in reading and mathematics. Others, like Kaplan, focus on test preparation. In any event, these firms are selling K–12 educational services to the public. There are no state licensing requirements for teachers in these firms (or for the firms themselves), and all indications are that this market is expanding.

Finally, there is a thriving private K–12 school system in the United States that long predates the public school system. Private schools routinely hire unlicensed teachers. Figure 2 provides some data on certification rates of private school teachers. The dependent variable is whether the teacher holds regular or provisional state certification in her primary teaching area: traditional public, private, and charter schools. Source: 1999–00 Schools and Staffing Surveys.
teaching area. The rate for the public sector is 89.8%, whereas the rate for private schools is much lower, particularly in nonreligious schools where just 48.8% of teachers are certified. The rates are lower still at the secondary level. In nonreligious secondary schools the certification rate is just 35.1%. Therefore, although private schools do hire certified teachers, they also hire substantial numbers of noncertified teachers as well. It should also be noted that charter schools, too, hire large numbers of uncertified teachers.

How does the academic quality of the uncertified teachers compare to that of the certified? One measure of teacher quality is the selectivity of the college from which the teacher graduated. Several production function studies find that the selectivity of a teacher’s undergraduate college is correlated with student academic achievement (Ehrenberg & Brewer, 1993, 1994; Summers & Wolfe, 1977; Winkler, 1975). The data in Table 2 suggest that private schools use this flexibility to trade off teacher certification to get higher academic quality for teachers. The share of teachers graduating from selective institutions, math and science majors, and academic majors is consistently higher in the noncertified population. A similar pattern is seen in charter schools. In other words, in terms of Figure 2, charter and private schools benefit from their ability to hire outside the certified candidate pool when an attractive noncertified applicant appears in the applicant pool.

**State Teacher Licensing Systems Are so Complex That No One Is in Compliance Anyway**

Proponents of raising the bar for teaching licenses assume that such proposals are feasible. However, I find that state licensing systems are already so complex that virtually no school district is in compliance anyway. This raises a serious question as to what we accomplish by raising bars.

Like all other states with which I am familiar, the state of Missouri issues a single license to practice medicine, law, dentistry, accounting, nursing, and veterinary medicine. However, in the area of K–12 education, the Missouri Department of Elementary and Secondary Education currently issues 260 certificates and endorsements (171 vocational, 89 nonvocational). However, that is only part of the story. There are levels of certification (permanent, provisional) for all of these and a host of “grandmothered” codes. As a consequence, there are 781 valid certification codes in the master teacher certification file. There is nothing unique about Missouri. Most other states have equally Byzantine systems for teacher licensing.

Further evidence on differences in personnel policies between traditional public and private and charter schools is found in Ballou and Podgursky (1997) and Podgursky and Ballou (2001).
### Table 2

**Measures of Teacher Quality in Public and Private Schools**

<table>
<thead>
<tr>
<th></th>
<th>Public Charter</th>
<th></th>
<th>Private Religious</th>
<th></th>
<th>Private Nonreligious</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Traditional Public: All</td>
<td>Certified</td>
<td>Not Certified</td>
<td>Certified</td>
<td>Not Certified</td>
<td>Certified</td>
</tr>
<tr>
<td>College selectivity: Most competitive</td>
<td>0.9</td>
<td>1.4</td>
<td>4.0</td>
<td>0.8</td>
<td>3.1</td>
<td>3.4</td>
</tr>
<tr>
<td>Other selective</td>
<td>22.1</td>
<td>26.7</td>
<td>29.5</td>
<td>18.5</td>
<td>21.4</td>
<td>30.3</td>
</tr>
<tr>
<td>Total selective</td>
<td>23.0</td>
<td>28.1</td>
<td>33.5</td>
<td>19.3</td>
<td>24.5</td>
<td>33.7</td>
</tr>
<tr>
<td>Math and science majors</td>
<td>10.1</td>
<td>10.1</td>
<td>12.8</td>
<td>11.1</td>
<td>10.8</td>
<td>11.6</td>
</tr>
<tr>
<td>Academic majors</td>
<td>33.7</td>
<td>40.9</td>
<td>56.3</td>
<td>37.2</td>
<td>55.5</td>
<td>39.0</td>
</tr>
</tbody>
</table>

*Note.* Data from 1999–00 Schools and Staffing Surveys.
How is it that the public is protected by a single license in other professions, yet K–12 education requires over 100? Is teaching a more complex endeavor? I believe that part of the answer is that in these other professions licensing is simply used to screen out incompetent practitioners but is not meant to control how labor is utilized in that sector. After a practitioner enters the profession, he is free to specialize in any field he chooses. Most doctors do proceed to earn certification in 1 of the 24 medical specialties, but there is no state requirement that they do so. If a medical clinic chooses to use a neurosurgeon to treat walk-in family practice patients, there are no legal impediments to doing so. Once licensed, lawyers are free to practice any type of law they choose. One does not read about a crisis of lawyers “practicing law out of field,” nurses “nursing out of field,” or dentists engaged in “dentistry out of field.”

In K–12 education, state regulators attempt to use the licensing system to control how teacher labor is allocated. The presumption is that local schools cannot be trusted to staff courses appropriately. Therefore, this complicated licensing system is the state’s clumsy attempt to monitor the performance of local administrators. In Missouri, school districts are routinely audited to determine whether the hundreds of different types of courses taught match to the right certificate or endorsements for the teacher of record. If states issued a single license in teaching as in other professions, most of the out-of-field teaching that is the subject of so much hand wringing would disappear.

The excessively complex licensing system in K–12 education can also be seen as a means by which teacher unions and schools of education engage in “rent capture” (i.e., using government regulation to produce private pecuniary gains). To make the case for higher pay and benefits for their members, any type of supply restriction is desirable from the point of view of teacher unions, so long as the added restrictions apply to new entrants and not dues-paying incumbents. However, high standards for program entry, as in medicine, would invariably drive many schools of education out of business. Therefore, a compromise between the teacher unions and the schools of education is to proliferate certificate areas, generating more demand for education school courses but also restricting supply to school districts.

As a consequence of the complex licensing systems that states have constructed, virtually no school district in the United States is in full compliance. The complexity of the state licensing systems make national tabulations of unlicensed, uncertified, or substandard certification difficult. Therefore, I illustrate this point with administrative data from two states. Figure 3 presents data for Missouri public K–12 school districts (I have excluded K–8 districts). On the vertical axis we measure the percentage of courses taught by teachers with inappropriate licenses during the 2001–02
school year. On the horizontal axis we measure spending per student in average daily attendance. Of 447 K–12 school districts, not a single district had no courses taught by an inappropriately licensed teacher (the average was 9.5%). Moreover, the prevalence of inappropriate licensed practice seems to have little to do with school revenues. The correlation between the rate of unlicensed teaching and spending per student is positive and statistically significant (.27).

My second example is decidedly nonrandom. Westchester County, New York, is home to some of the wealthiest households and highest paid schoolteachers in the United States. The schools in these exclusive communities are appropriately compared to the very best private day schools. In Figure 4, I have plotted percentage of uncertified teachers and median teacher salaries for school year 2000–01 from the most recent report of the New York State Department of Education. *(Uncertified* means the percentage of classroom teachers who teach more than 20% of their time in a subject or subjects for which they hold no certification or emergency license.) Despite very high salaries, no district in Westchester County has fewer than 2% of their teachers uncertified. Note that Scarsdale, which boasted a 2000–01 median teacher salary of $90,191, had 6% of its teachers uncertified. (As a regular reader of the *New York Times*, I have yet to read about Scarsdale parents complaining about the quality of their uncertified teachers.) If not a single school district in what may be the highest spending county in the United States is in full compliance with the New York State

Figure 3. Percentage of courses taught by teachers with inappropriate or no licenses by expenditure per pupil in average daily attendance: Missouri K–12 public school districts, 2001–02. Source: Missouri Department of Elementary and Secondary Education.
law, this raises serious questions as to how we can contemplate raising the bar on teacher licensing.

The Model

In light of the six propositions laid out in the previous section, the case for a more flexible licensing model is relatively straightforward. State education regulators should protect the public by focusing on what they can measure (student learning) and not on what they cannot (teacher quality). That is, they should focus on education outputs rather than education inputs. As noted earlier, research to date suggests that teacher quality as measured by student learning is idiosyncratic and not well measured by anything that state regulators are in a position to monitor. Local school administrators, on the other hand, are in a good position to monitor teacher classroom performance. Therefore, the model that emerges has several features.13

The first, and most important, is getting incentives right. Local administrators must be held accountable for student learning, and state regula-

13The model proposed here is similar to that in Hess (2001), which he described as “competitive certification.”
tors need to focus their attention on monitoring student learning. Nearly all states have developed standards for what students should be learning at various grade levels and assessments of the learning that are actually occurring by grade in schools and districts. These data are now routinely provided to parents and to the public at large. Schools that demonstrate persistently poor performance in terms of student learning increasingly face administrative sanctions. Another important mechanism for producing accountability is school choice. Indeed, one of the most important protections for parents against incompetent teachers is to give them the option to choose another classroom or school if their assigned teacher fails.

If a school accountability regime is in place with information on performance widely available, with state monitoring of school performance, and with parental choice of schools, the role of the state in monitoring “teacher quality”—something that they cannot measure anyway—should whither away. As far as teacher licensing is concerned, the guiding principal for the state should be a simple “do no harm” standard. Certainly, teachers should undergo a careful criminal background check. Current bureaucratic impediments to removing teachers who have been convicted or indicted for serious criminal offenses should be removed. It is reasonable that teachers should also be required to hold a bachelor’s degree.

Tests of general academic knowledge and subject matter knowledge are reasonable. However, as noted earlier, it is likely to be counterproductive for states to set high cut scores for these exams. A more attractive approach (which is likely to survive legal challenge) is to set relatively low cut scores and provide information on the candidate’s scores to the school districts. From an economic point of view, the current system makes little sense. Teacher candidates may spend hundreds of dollars taking licensing exams. The testing companies then take these continuous test scores and collapse them into a “pass” or “fail” grade. That is the only result that school districts ever see. Large testing companies such as the Educational Testing Service will not provide these exam scores to school districts. Indeed, they make the tenuous argument that the scores only have “validity” for licensing but not for hiring. Imagine if colleges or professional schools received similar information. Continuous SAT or GRE scores would be collapsed into binary pass or fail measures indicating that a student was “good enough for graduate study” or “not good enough for graduate study.” The test scores would have “validity” only for determining who can enter the market for graduate study but not for the admissions decisions of any graduate program. Academics would immediately dismiss such a proposal as absurd; yet, somehow this type of reasoning holds sway in occupational licensing.
Of course, providing flexibility for schools to audition many candidates means that schools must have the ability to act on that information. In fact, in most states school districts have considerable leeway to not rehire teachers during an initial probationary period of 2 to 5 years. After that, once teachers are “tenured,” or enjoy the right of automatic contract renewal, it becomes very costly to dismiss teachers for anything but the most negligent job performance. Sensible reforms would make it easier to not renew the contracts of ineffective senior teachers.

Even if a “magic bullet” in terms of teacher training or testing were found, it would be many years before new, more effective teachers diffuse through the teaching workforce. Teacher turnover averages roughly 6% to 8% annually, and about one half of new teacher hires are returning teachers or interdistrict transfers (Broughman & Rollefson, 2000). Therefore, only a very small share of teachers are affected by licensing reform in any year. Schools have information on the job performance of 90% of job candidates who do not turn over. They need the flexibility and incentives to use this performance information in personnel matters, including pay and contract renewal.

To summarize, the most efficient flexible regime would have the following features:

- Accountability for student learning through testing, sanctions, and parental choice.
- State regulators who actively promote a competitive market in teacher quality and protect schools from anticompetitive practices on the part of teacher unions, schools of education, or other education producer organizations.
- Minimal state licensing standards for teachers: criminal background check, bachelor’s degree, test of general and content knowledge.
- Full information on teacher test results provided to school administrators.
- Award of a permanent or full license on the basis of successful job performance.

Statutory Changes

State boards or professional certification boards generally have considerable discretion as to the content and requirements for teaching licenses. There is certainly ample precedent for a flexible policy. Many states have now created alternative routes to teacher certification that provide a good model for what I have proposed. The Alternative Certification Program in Texas and the Intern Program in California are examples. In such pro-
grams, prospective teachers must hold a bachelor of arts degree, pass exams, and demonstrate content knowledge to receive a probationary license. After a probationary period of 2 to 3 years (which in these states includes on-site or off-site professional development), mentored practice, and satisfactory evaluations by supervisors, the teachers receive a standard license (Feistritzer, 2003). Such alternate route teachers satisfy the highly qualified teacher requirement of the No Child Left Behind Act.

Role of Education Schools and Organizations

Some years ago, Myron Lieberman (1993) perceptively described the K–12 education industry as “producer dominated.” I find this description accurate, particularly in the area of teacher training and licensing where education school faculty and teacher organizations dominate the regulatory process. Both the teacher unions and the education schools have a common interest in “professionalizing” teaching by restricting supply through proliferation of certificates and suppressing competition in teacher training (e.g., by preventing entry by new institutions). Moreover, I also find a similar view in state education departments, which embrace “teacher professionalization” as part of their mission along with raising student achievement.

Moving to the flexible model described earlier would create strong competitive pressures for teacher training organizations to improve. Simply put, if education school courses are no longer required to hold a teaching license, then the monopoly power of schools of education largely disappears. If the pedagogical training offered by schools of education does, in fact, raise student achievement, then graduates from such programs will enjoy a competitive edge in the labor market and have more desirable job offers than their untrained peers. In that case, students will flock to such programs. On the other hand, if a teacher training program cannot attract adequate enrollments in a market in which job candidates and employers have flexibility as to the credentials, then it will go out of business. Effective programs will thrive and ineffective programs will whither away.

Any institution, public or private, would be free to enter the market and provide teacher training. If a history department at a small liberal arts college wanted to train teachers by offering a course preparing majors to teach in secondary schools and worked out student teaching arrangements with nearby schools, they would be free to do so. Unlike the current system, they would not be required to “partner” with a school of education.

When a major bank comes to a university campus to recruit candidates for management positions, they can interview finance majors in the business school, economics majors in the College of Arts and Sciences, or oper-
There are no licensing entry barriers creating a monopoly for a particular college. Opening up the teacher training market to competition is a more effective way to “fix” the quality problem in schools of education than regulation by state departments of education or mandatory accreditation by NCATE. It is common knowledge that low-quality schools of education have succeeded in securing approval from state education agencies as well as accreditation from NCATE. Clearly, both of these quality control systems have failed to weed out programs with academically mediocre students or low pass rates on state licensing examinations.

Finally, a flexible regime will require some change in gestalt among the regulators in state education agencies. The primary objective of state regulators should be increasing student learning and narrowing achievement gaps. How schools and districts do this—assuming the behavior is ethical and legal—should be of secondary importance. Children are protected from incompetent practice by monitoring learning through regular testing and by school choice.

On the other hand, state education agencies should not be in the business of promoting teacher professionalism. Here is where a change of gestalt is required. Promoting teacher professionalism is a role for private organizations. If organizations like teacher unions and (private or public) schools of education choose to promote teacher professionalization by securing accreditation of teacher training programs by NCATE, that is their choice. However, state education agencies should not impose these choices on the entire market through manipulation of licensing requirements.

Moreover, it should be recognized that in some educational endeavors, the most cost-efficient way to promote student achievement might involve de-professionalization of teaching. In some schools or programs, highly scripted curricula, distance learning, or computer-based instructional programs may reduce the need for highly trained teachers. Schools should face incentives to adopt the most cost-efficient approaches to promoting student learning, whether they advance teacher professionalization. If in the name of teacher professionalization state education agencies encourage schools to spend additional resources when more cost-effective means for instructional delivery are available, then resources that might have been used to lower class sizes or otherwise enhance student welfare have been diverted. In fact, it may be the case that the resources saved on teachers might more effectively promote student learning outside of K–12 schools (e.g., in better medical care for poor women, reduction of crime and drugs in low-income communities, or preschool care).

Should states continue to regulate teacher training programs? Because there is virtually no reliable research establishing a causal link between
any program of teacher training and student achievement gains, state reg-
ulators have little basis for regulating teacher training programs. Some
states and federal legislation (Title 2 of the Higher Education Act, P.L.
105–244, 1998) have focused on pass rates on teacher licensing exams as
one criterion to judge programs. In my own analysis of Missouri data, I
have found that the most important predictor of whether a teacher passes a
Praxis II exam is his or her American College Test (ACT) score. Once we
control for student ACT scores, there are few significant differences be-
tween institutions. In addition, there is wide dispersion of test scores
within any institution. If pass rates primarily reflect the academic ability of
the students entering the teacher training program, it makes little sense to
use pass rates to assess the quality of a training program.

Because there is no reliable research base for approving or denying a
teacher training program, and institutional pass rates are largely driven by
the quality of program entrants, a reasonable approach would be for regu-
lators to be fairly liberal in program approval. The primary mechanism to
raise program quality would be market pressures and not regulation. If a
training program does a poor job of preparing teachers to meet state edu-
cation standards, graduates from such programs will receive fewer job of-
fers in the market and enrollments will decline. Eventually, the program
may leave the market altogether. By the same token, programs that pro-
duce high-quality teachers will attract many applicants and expand.

In sum, state education agencies should create strong incentives for
schools and districts to raise student achievement and give them flexibility
as to how to they get the job done. If one or another model of professional-
ism promoted by private organizations is a cost-effective way for schools
to achieve this end, then professionalization will expand. If it does not,
then it will languish. However, this is not a matter of public policy.

Evaluation

Effective educational policy requires that educational interventions be
evaluated. We find ourselves in the current situation precisely because the
education research community has for decades failed to conduct research
on teacher quality that meets scientific research standards. However, relax-
ing licensing standards will generate nonexperimental data on the causal
relation between teacher credentials and student achievement. Ironically,
the current system, by encouraging homogeneity, reduces our ability to as-
ess teacher effects. If all teachers in a school district matriculate from the
same teacher-training program, then it is impossible to estimate the effect
of that program on student achievement. The best way to assess the effect
of a treatment variable on an outcome variable is to maximize variation of
the treatment variable. Relaxing entry barriers will generate much more natural variation in the workforce in the credentials and training of teachers. This will permit better evaluation of the effects of teacher credentials and training.

It is interesting to note that one factor that has often been ignored in the research literature on teacher certification has been the effect of the regulatory regime on the distribution of teaching certificates (i.e., What was the process that produced the observed distribution of teaching credentials?).

Consider the effect of emergency licenses or waivers. In the current regime, school districts are not supposed to hire such teachers if certified teachers are available. Therefore, the data we observe on teachers with emergency versus full licenses is generated by a process in which head-to-head competition between certified and noncertified teachers as depicted in Figure 2 is not permitted. It may be that in a more competitive regime, teachers with emergency licenses would be of higher quality. Why would this be? If schools were free to recruit emergency certified teachers in the same way that they recruit licensed teachers, many who do not currently pursue teaching jobs might be enticed to apply. The result would be a much larger pool of talented emergency candidates. With a larger pool from which to choose, those hired would presumably be of higher quality. As noted earlier, on average, schools will end up with a better hire if they have 50 applicants for a job than if they have 5.

What Could Go Wrong?

Does this approach involve risk? Yes, it does. Relaxing entry restrictions into teaching will permit greater flexibility for schools to seek out the best teachers and meet the performance targets set by state regulators. Regulators would protect parents and children against incompetent practice by monitoring student learning and making such data widely available. Parents would also be empowered to protect their children from poor teachers by giving them more choices among schools. Markets and competition are ultimately the best guarantor of quality in the provision of almost any service.

However, would such a system produce greater exposure of children to incompetent teachers? Ultimately, this is an empirical point; however, I do not believe that most objective observers of current licensing systems would argue that it is particularly effective in screening out incompetent practitioners. Indeed, as an empirical matter, it is likely the case the greatest harm from incompetent teacher practice comes not from novices but rather from experienced, licensed teachers who are protected by tenure statues.
Conclusion: The Tail Is Wagging the Dog

Policy debates about teacher quality have tended to dwell on teacher training and licensing. However, there is little research indicating that the types of licenses that teachers hold or the type of pedagogical training program they have passed through have a significant relation to student performance. However, even if effective changes in licensing or training were identified, it would be many years before significant effects on student achievement would be obtained. This is because the number of inexperienced teachers hired in any year is very small relative to the stock of incumbent or experienced teachers (Broughman & Rollefson, 2000). In other markets, the best we expect from licensing is to screen out incompetent new practitioners. However, the quality or performance for incumbents is primarily determined by incentives: Experienced dentists who do a poor job on our teeth lose customers; those who perform incompetently get sued.

I have argued that attempts to address the teacher quality problem by raising bars in teacher licensing are likely to make things worse rather than better. Because there is so little reliable research to guide setting criteria for market entry, and such modest effects of teacher credentials in the current research, all such approaches are likely to accomplish is a reduction in the size of the teacher applicant pool with little change in the average productivity of the applicants. However, in a world of uncertain teacher productivity, it is in the interest of school districts to have more candidates to audition than fewer.

A more productive approach is for state regulators to focus on what they can measure (student achievement) and not on what they cannot (teacher quality). State regulators should make sure local school administrators have adequate instructional resources and strong incentives for raising school performance. They should use licensing to reduce the likelihood that a demonstrably incompetent teacher is put into the classroom. A prudent standard in this regard is a test of general academic skills, and more specialized tests covering the teaching fields and subjects to which the teacher is assigned. However, the most important role for teacher licensing reform is permissive or enabling. We need to make sure that these procrustean licensing systems do not stand in the way of entrepreneurial school administrators who are responding to the incentives we create for improving student performance.

Rather than dwell on the credentials and training of the 3% to 4% newly minted teachers hired each year, it is much more important to create strong performance incentives for the other 95% of teachers. Performance incentives are absent when pay is set by rigid salary schedules and tenure sys-
tems that protect teachers whose poor performance warrants dismissal. Rather than expend further resources seeking indirect measures of job performance, such as licensing exam scores or teaching portfolios, it would be far more productive to make better use of available information on teaching performance for the 95% of incumbent teachers. Dismissing 2% of the least productive teachers in the workforce based on current job performance would surely have a much larger effect on student achievement than marginal changes in the training or licensing of 3% to 4% of newly minted hires. Finally, it is important to address the role that collective bargaining contracts play in stifling efforts to raise teacher quality, particularly in urban school districts.

Teacher quality and effort primarily is a management problem, not a licensing problem. School principals in major urban districts often lack the ability to select their teachers, dismiss ineffective teachers, and are often severely hampered by collective bargaining agreements (and licensing regulations) in how they can assign their teachers or staff vacant positions. With district-wide salary schedules, school administrators have a passive role in setting individual teacher pay. The teacher quality policy debate needs to focus on creating efficient incentive structures and reducing constraints on effective management. Teacher licensing is of secondary importance.

References


