A Qualified Teacher in Every Classroom?
Appraising Old Answers and New Ideas

Edited by
Frederick M. Hess,
Andrew J. Rotherham,
and Kate Walsh

HARVARD EDUCATION PRESS
2004
Model 4
Improving Academic Performance in U.S. Public Schools
Why Teacher Licensing Is (Almost) Irrelevant
Michael Podgursky

INTRODUCTION: SIX PROPOSITIONS ABOUT TEACHER LABOR MARKETS
The "flexible" regime would call for state education agencies to relax their requirement that teaching candidates graduate from approved teacher-training programs. It is flexible concerning testing, in that it would set relatively low testing requirements for teachers as well. Teachers might need to pass tests of general knowledge and content in areas in which they specialize, but the bar would not be particularly high. Tests would screen out the academically incompetent, and criminal background checks would screen out potentially dangerous teachers. However, the primary focus of state regulation would be on monitoring student learning and educating the public about state education standards and school performance, and not regulating teacher training and licensing. Rather than rely solely, or even primarily, on teacher licensing, state monitoring of student learning and school choice would play a larger role in protecting parents against incompetent teachers. Private associations of teachers and teacher educators would be free to promote whatever models of teacher "professionalization" or accreditation they choose, and districts would be free to hire these candidates as
they choose; however, state regulators would not use teacher licensing to impose any particular model of professional training.

The flexible option exists alongside other reform models that are discussed in the companion papers in this book. The alternative models set forth—the "professional," "portfolio," and "candidate-centered" models—all restrict the supply of teacher programs in ways that are not likely to produce any significant increase in overall teacher quality.

To lay the groundwork for a discussion of the flexible regime, I begin with some observations concerning reform on teachers and on the economics of licensing.

1. 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 on teacher licensing, they should have solid evidence that the criteria used to exclude teacher candidates from the market have a demonstrable relationship 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 (including Daniel Goldhaber in this volume) find it is very thin.

Such a claim 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. However, there is widespread consensus in the social-science research community that scientific evaluation of social-policy programs (including education) requires either randomized experimental study design or non-experimental longitudinal data on participants. Unfortunately, the little research on teacher testing or licensing that meets either standard is tentative and inconclusive.

Randomized experimental design is the "gold standard" for social-policy research. In the context of 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 (in, e.g., experience) teachers within a school. Unfortunately, no existing research on teacher credentials or training meets this standard, although the Institute for Education Sciences of the U.S. Department of Education is promoting such studies (Moesteller & Schuch, 2002; U.S. Department of Education, n.d.), and some are under way. Thanks to these efforts, it is likely that we will have experimental evidence on teacher licensing and training five years from now.

If randomization is not feasible, and often it is not, then one must rely on non-experimental 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 test 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 one 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 (e.g., Anderson, Barnow, & Sanders, 2003; Goldhaber & Brewer, 1997; Rivkin, Hanushek, & Kain, 2000; Sanders & Horn, 1994).

Studies that are not rigorously designed (i.e., do not randomize or control 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 background of students, which 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." 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 socioeconomic status (SES) students. Because SES has a 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.

The number of studies of teacher certification that meet the minimum methodological standards outlined above 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 socioeconomic status. The results of these studies (both by Goldhaber and Brewer, both using the National Longitudinal Educational...
Survey of 1998) were mixed. 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 Hamusheke and Rikvin (2003) focusing on "high-quality" studies that meet the standards described above 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 effect. Of the three finding a significant effect, two were positive and one was negative.5

In short, the research foundation for "raising the bar" with teacher tests or raising standards for schools of education is very 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, while usually statistically significant, are generally modest in effect.

2. Teacher Effects on Student Achievement Are Quantitatively Important but Idiosyncratic

Does this mean teachers do not matter? On the contrary, while the effect of measured teacher characteristics is small, one consistent finding is that there seems to be considerable variation in teacher effectiveness between classrooms. Thus, if comparing the effect on student learning of the top and bottom 20 percent of teachers ranked by classroom value added, 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 (e.g., experience, or graduate degrees, a result highlighted in a survey by Golish et al. (2002), Hamusheke and Rikvin (2003), summarizing their own and others' research, come to the same conclusion.6

The growing value-added 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, but that traditional measures of teacher quality—experience, master's degrees, education coursework—explain virtually none of the variation in teacher effectiveness.7

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

Even if the evidence for teacher licensing, testing, or a particular program of pedagogical training is weak, a skeptic might say, Why not raise the entry bar anyway, on the chance that it might work? Many public policies are enacted based 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 cost free, then one might make the case for their implementation on the chance that some benefits would accrue. However, they are not cost free, and there is a 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, etc. Annual costs per student in higher education currently are roughly $27,000 per year, although students on average only pay part of this cost. Tests are less costly, but the fixed costs of developing 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 that teaching candidates take a year and a half of teaching courses (including student teaching), this is a costly investment. Even at the minimum wage, this amounts to more than $15,000. One perverse result of conditioining labor-market entry on "seat time" in pedagogy courses is that candidates with greater academic skills, who presumably can attract a higher alternative wage, face higher costs in securing a teaching license (Ballou & Podgursky, 1997). These direct costs are resources that might have been put to better use in improving schools or the lives of poor students.

Only about 60 percent of all teacher candidates graduate from teacher-training programs recognized by the National Council for the Accreditation of Teacher Education (NCATE). Closing all teacher-training programs that do not secure NCATE accreditation, as proposed by the National Education Association and the National Commission on Teaching and America's Future (NCTAF), would almost surely restrict the flow of newly trained teaching candidates. So, too, would raising the cutoff score on teacher-licensing exams. If the cutoff is raised from the 20th to the 30th percentile on the elementary education exam, then 10 percent of potential applicants are excluded from the applicant pool.

A common feature of all these policies is that they would reduce the pool of applicants 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 their
A QUALIFIED TEACHER IN EVERY CLASSROOM?

607x1178

FIGURE 1 Overlapping Ability When the Average Certified Applicant Is Better than 60 Percent of Uncertified Applicants

average applicant, it is better to have more applicants. This commonsense point is borne out in many studies conducted. Other things being equal, the average quality of graduate students will be greater in a program with 200 applicants than one with twenty, even if the average quality of applicants is the same in the two cases. The expected return from twenty lottery tickets is higher than from one.

The example above assumes that we only sample from the certified, "high-quality" applicant pool. Can a district similarly benefit from access to a pool of lower average quality uncertified applicants? The answer is yes. Suppose that five uncertified applicants randomly drawn from the distribution in Figure 1 are allowed to apply for the job, along with five certified applicants. Even though the uncertified applicants are of lower average quality than the certified candidates, 57 percent of the time the certified candidate is better. With two 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 five random applicants from the certified population, but no uncertified applicants. 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 below), a larger applicant pool means better quality hires. Because you are hiring the best applicant from a pool, not the

concentration 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. Because they are prevented from considering any unlicensed applicants, school districts are forced to hire the worst certified candidate even if a superior uncertified candidate is available. "Raising the bar" shrinks the size of the applicant pool. The new pool is better in terms of whatever the regulators specify (e.g., more NCAE graduates, higher Praxis I scores), but school administrators have less ability to select among candidates based on 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. Though these data are hypothetical, I believe they represent the picture emerging in the "teacher effects" literature, namely, that the individual variation in the classroom performance of a teacher is large relative to any measurable teacher characteristic, such as certification. As I indicated above, the evidence concerning teacher certification is mixed at best. However, for the sake of argument, I have assumed a positive effect: the average certified teacher is better than 60 percent of uncertified teachers. Based on the review of the research literature discussed above, I see this as an upper bound estimate for certification effects. However, the conclusions that follow do not hinge on this assumption. If, for example, we assumed that the average certified teacher is better than 80 percent of uncertified teachers, our basic conclusions would not 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 teachers, that the school district has a single vacancy, and that it is free to hire the best candidate, certified or not. Imagine that a single candidate applies at random from each pool (certified and uncertified). What is the probability that the certified candidate is the superior teacher? It turns out that 57 percent of the time the certified candidate is better. With two 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 five random applicants from the certified population, but no uncertified applicants. 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 below), a larger applicant pool means better quality hires. Because you are hiring the best applicant from a pool, not the
4. 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?" I believe the argument implied by the question is this: While it may be true that no rigorous evidence exists 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 paper of its own; however, I will briefly explain two important reasons the analogy is inappropriate.

There is a deep body of scientific research in medicine. Commitment to scientific research methods pervades medical schools, 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 about the quality of medical services they are buying. Because of the complexity of the knowledge base in medicine and the high cost of mistakes, it is relatively easy to argue that government licensing is required to screen out incompetent practitioners and protect consumers.

Does this model apply to 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.

Even at leading research universities, most 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 like those used 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 don't do is largely irrelevant since they train relatively few of the nation's classroom teachers. The primary suppliers of classroom teachers (as opposed to doctors) are state colleges, most of which were once teachers' colleges, where 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, 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 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 (practicing teachers) are not trained to evaluate scientific research. I would venture that most practitioners and education school professors in these fields would not even view the scientific method as the most useful method of inquiry in their field. (I do not believe that the medical model is appropriate for departments of history or English, for similar reasons.)

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. While the judgment of English, mathematics, and elementary school teachers as to the best ways of teaching a subject certainly deserves respect and defense, there is little evidence to suggest that parents cannot make informed choices among practitioners who approach their craft differently. This leads us to the next proposition.

5. Unregulated Markets in Education and Training Work Well

The case for the medical analogy would be strengthened if there were pervasiveness of "market failure" in unregulated markets for education and training. While I am not aware of widespread unlicensed practice of surgery, unlicensed training and schooling is pervasive in our economy. Indeed it is the norm rather than the exception. These markets seem to work quite well
with little or no government regulation. 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 billion and $43 billion (in 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 armed services have taken millions of high school dropouts and graduates and given them high-quality training in technical and specialty fields, along with basic literacy and numeracy skills, turning millions of young men and women with limited elementary and secondary education into trained aircraft mechanics, radio operators, supply clerks, and the like. Nearly all of this was accomplished by unlicensed instructors.

Approximately six million students are enrolled in two-year community colleges. Much of the coursework offered in these community colleges is remedial, covering 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 or successful transition to four-year baccalaureate institutions, these community colleges are successfully delivering K-12 educational services.

Many students receive K-12 educational services from private tutoring firms, which 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. The rate for the public sector is 89.8 percent, whereas the rate for private schools is much lower, particularly in nonreligious schools, where just 48.8 percent of teachers are certified. The rates are lower still at the secondary level. In nonreligious secondary schools the certification rate is just 55.1 percent. Thus, while private schools do hire certified teachers, they also hire substantial numbers of uncertified teachers. 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 certified teachers? 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). Dale Ballos and I have shown 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 uncertified population (Ballos & Podgursky, 1997). In more recent work, we find a similar pattern in charter schools. In other words, in terms of Figure 1, charter and private schools benefit from their ability to hire from beyond the certified candidate pool when an attractive uncertified applicant appears.
6. State Teacher Licensing Systems Are So Complex That No One Is in Compliance

Propositions 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 of what we accomplish by raising bars.

Like all other states with which I am familiar, Missouri issues a single license to practice medicine, law, dentistry, accounting, nursing, and veterinary medicine. However, in 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 751 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.

How is it that the public is protected by a single license in other professions, yet K-12 education requires more than 1000 is teaching a more complex endeavor? I believe part of the answer is that, in the other professions, licensing is used simply to screen out incompetent practitioners, not to control how labor is utilized in the sector. After a practitioner enters the profession, he is free to specialize in any field he chooses. Most doctors do proceed to earn certifications in one of the twenty-four recognized medical specialties, but there is no state requirement that they do so. If a medical clinic wants to use a neurosurgeon to treat walk-in family practice patients, it can. 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," "nursing out of field," "dentistry out of field," etc. States issued a single license in teaching as in other professions, the "out of field" teaching that is the subject of so much hand-wringing would disappear.

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. Thus, the 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 they offer match the certificate or endorsements for the teacher of record.

The excessively complex licensing system in K-12 education can also be seen as a tool for what economists call "rent capture" (using government regulation to produce private pecuniary gains) by teachers unions and schools of education. 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 teachers unions, so long as the added restrictions apply to new entrants and do not affect dues-paying incumbents. However, high standards for program entry, as in medicine, would invariably drive many schools of education out of business. Thus, teachers unions and schools of education benefit from the proliferation of certificate areas, generating more demand for education school courses but also restricting supply to school districts.

As a consequence of the complex state licensing systems, virtually no school district in the United States is in full compliance. The complexity of the state licensing systems makes national tabulations of unlicensed, uncertified, or substandard certification difficult. Thus, I will 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 in a district taught by teachers with inappropriate licenses during the 2001-2002 school year. On the ho
in Scarsdale, which boasted a 2000-2001 median teacher salary of $90,191, 6 percent of teachers were 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 one school district in what may be the highest-spending county in the United States is in compliance with the New York law, this raises serious questions about how we can seriously contemplate raising the bar and further restricting supply.

**THE MODEL**

In light of the six propositions laid out above, the case for the "flexible" model is relatively straightforward. State education regulators should protect the public by focusing on what they can measure (student learning, or "education outputs"), and not on what they cannot (teacher quality, or "education imports"). As noted above, research 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. Thus, the model that emerges has several features. The first, and by far most important, is getting incentives right. Local administrators must be held accountable for student learning, and state regulators 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 is 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 student learning performance 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 an accountability regime is in place, with information on performance widely available, state monitoring of school performance, and parental choice of schools, then the role of the state in monitoring "teacher quality" (something they can't measure anyway) should wither away. The guiding principal for the state in issuing teacher licenses should be simply, "Do no harm." Certainly teachers should undergo a careful criminal background check. Bureaucratic impediments to removing teachers who have been convicted or indicted for serious criminal offenses should be removed. It is also reasonable to require that teachers hold a bachelor's degree.
Tests of general academic knowledge and subject-matter knowledge are reasonable. However, as noted above, it is likely to be counterproductive for states to set high cut-scores for these exams. A more attractive approach for states is to set relatively low cut scores (which is likely to survive legal challenge) 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 spend hundreds of dollars taking licensing exams. Instead, they could be used to reduce teacher turnover in school districts.

Teacher candidates spend hundreds of dollars taking licensing exams. Instead, they could be used to reduce teacher turnover in school districts.

The testing companies then take these test scores and collapse them into a "pass" or "fail" grade. Many states now have 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 programs in California are examples. In each program, prospective teachers must hold a bachelor's degree, pass specific content knowledge tests, and have a year of experience teaching in a classroom. The test scores would have "validity" only for low-income students. The test scores would not do the job.

Of course, providing flexibility for schools to audition many candidates in most states, school districts have considerable leeway not to rehire teachers during an initial probationary period of two to five years. After that, the contracts of ineffective teachers would be renewed.

Even if a "magic bullet" for teacher training or testing were found, it would be decades before new, more effective teachers would have sufficed to improve the quality of teaching. To summarize, the most efficient flexible regime would have: 1) accountability for student learning through testing, sanctions, and parental choice; 2) state regulators who actively promote a competitive market in teacher quality and protect schools from anticompetitive practices on the part of teacher unions, states, and education producer organizations; and 3) minimal state licensing standards for teachers (criminal background check, bachelor's degree, test of general and content knowledge) and a full 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 requirements for teaching licenses. There is ample precedent for a flexible policy. Many states now have 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 programs in California are examples. In each program, prospective teachers must hold a bachelor's degree, pass specific content knowledge tests, and have a year of experience teaching in a classroom. The test scores would have "validity" only for low-income students. The test scores would not do the job.

Of course, providing flexibility for schools to audition many candidates in most states, school districts have considerable leeway not to rehire teachers during an initial probationary period of two to five years. After that, the contracts of ineffective teachers would be renewed.

Even if a "magic bullet" for teacher training or testing were found, it would be decades before new, more effective teachers would have sufficed to improve the quality of teaching. To summarize, the most efficient flexible regime would have: 1) accountability for student learning through testing, sanctions, and parental choice; 2) state regulators who actively promote a competitive market in teacher quality and protect schools from anticompetitive practices on the part of teacher unions, states, and education producer organizations; and 3) minimal state licensing standards for teachers (criminal background check, bachelor's degree, test of general and content knowledge) and a full award of a permanent or full license on the basis of successful job performance.

**The Role of Education Schools and Organizations**

Some years ago, Myron Lieberman (1994) 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 of new institutions). 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 above would create strong competitive pressures on other 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 jobs 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, then it will go out of business. Effective programs will thrive; ineffective programs will wither away.

Any institution, public or private, would be free to enter the market and provide teacher professionalization. State education agencies encourage schools to spend additional resources when more cost-effective means for instructional delivery are available, then resources may be diverted from other goals like lowering class size or otherwise enhancing student welfare. In fact, it may be that resources saved on teachers might more effectively promote student learning outside of K-12 schools, for example, 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? Since there is virtually no reliable research establishing a causal link between any program of teacher training and student achievement gains, state regulators have little basis for regulating teacher-training programs. Some state and federal legislation (Title 2 of the Higher Education Act) have focused on pass rates on teacher licensing exams as a 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 ACT score. Once we control for student ACT scores, these are few significant differences between institutions. In addition, there is wide dispersion of test scores within any institution.

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 to regul-
tors to be fairly liberal in program approval. The primary mechanism for raising program quality would be market pressures, not regulation. If a training program does a poor job of preparing teachers to meet state education standards, graduates from such programs will receive fewer job offers in the market and enrollments will decline. Eventually the program may leave the market altogether. By the same token, programs that produce high-quality teachers will attract many applicants and expand. Protection against demonstrably incompetent teachers would come from testing individual tidbits, not screening programs.

In sum, state education agencies should create strong incentives for schools and districts to raise student achievement and give them flexibility as to how they get the job done. If one or another model of professionalism 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 for 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. Relaxed licensing standards will help generate non-experimental data on the causal relationship between teacher credentials and student achievement. Ironically, the current system, by encouraging homogeneity, reduces our ability to assess teacher effects. If all teachers in a school district matriculate from the same teacher-training program or hold the same credentials, then it is impossible to estimate the effect of that program or credential 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 a much more natural variation in the credentials and training of teachers in the work force. This will permit better evaluation of the effects of credentials. Perhaps alternatively certified school administrators will better appreciate the importance of random assignment experiments in schools and help move that project along as well. Certainly charter school operators have shown a willingness to experiment with new approaches in personnel policy and educational practice.

Interestingly, 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. That is, 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 uncertified teachers if certified teachers are available. Thus, 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 uncertified teachers, as depicted in Figure 1, is not permitted. It may be that 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 don't currently pursue teaching jobs might be enticed to apply. The result would be a much larger pool of talented emergency candidates. Those hired from a larger applicant pool would presumably be of higher quality. As noted above, on average schools would end up with a better hire if they have fifty applicants for a job than if they have five.

**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. Regulations 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 greater choices among schools. Markets and competition are ultimately the best guarantee 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 the current licensing system would argue that it is particularly effective in screening out incompetent practitioners. Indeed, as an empirical matter, it is likely that the greatest harm from incompetent teachers comes from experienced, licensed teachers who are protected by tenure status, not from novices.

**CONCLUSION: THE TAIL IS WAGGING THE DOG**

Policy debates about "teacher quality" have tended to dwell on teacher training and licensing. Yet, there is little research indicating that the types of licenses that teachers hold or the type of pedagogical training program they have passed through has a significant relationship 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 obtain. This is because the number of inexperienced teachers hired in any year is very small relative to the stock of incumbent or experienced teachers. In other markets, the best we expect from licensing is to screen out incompetent new practitioners. However, the quality of performance for incumbents is primarily determined by incentives: experienced dentists who do a poor job 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. All such an approach is likely to do is reduce the size of the applicant pool with little change in the average productivity of the applicants. In a world of uncertain teacher productivity, it is in the interest of school districts to have more candidates to audition rather than fewer.

A more productive approach is for state regulators to focus on what they can measure (student achievement), not on what they can't (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 material 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 those Fee for Service licensing systems do not stand in the way of school administrators who are responding to the incentives we are creating for improved student performance.

Rather than dwell on the credentials and training of the 3 to 4 percent of teachers who are new each year, it is much more important to create strong performance incentives for the other 95 percent of teachers. Performance incentives are absent when pay is set by rigid salary schedules and tenure systems that protect teachers whose poor performance warrants dismissal.

Rather than expend further resources seeking indirect measures of job performance, like licensing exam scores or teaching portfolios, it would be far more productive to use available information on teaching performance for the 95 percent of incumbent teachers. Dismissing the least productive 2 percent of teachers based on current job performance is likely to have a much larger effect on student achievement than marginal changes in the training or licensing of the 3 to 4 percent of teachers who are newly hired. Finally, it is important to address the role that collective bargaining contracts play in stalling efforts to raise teacher quality, particularly in urban school districts.

I would like to thank Youn Soel and Erin Allen for research assistance, and Michael Wolkoff and conference participants for thoughtful comments.
CHAPTER 10
Model 4: Improving Academic Performance in U.S. Public Schools

Michael Podgursky

1. For example, a 1996 report of the National Commission on Teaching and America's Future, a self-appointed commission including the president of National Council for the Accreditation of Teacher Education (NCATE) and the National Board for Professional Teaching Standards (NBPTS) noted: "Although hundreds of studies have shown that teachers... are more effective than those who are unqualified, the practice of hiring untrained teachers continues... 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" (National Commission on Teaching and America's Future, 1996, pp. 11, 23-24, 15).

2. A third model, most frequently used by economists, is the "natural experiment" ( Heckman, Lalonde, and Smith, 1999). This has not been widely employed in the teacher literature. An exception is Jacob and Ingberman (2002), who use longitudinal student-level achievement data for Chicago public school students and exploit a quirk in the administrative regulations to create a " quasi-experiments " to examine the effect of teacher training on student achievement. However, even with " natural experiments " longitudinal data is highly desirable.

3. A recent study by Hooyse (2001) highlights the importance of three socioeconomic variables and their potential for producing bias in studies of teachers and student achievement. Hooyse analyzed the effect of family, neighborhood, and school input variables on student achievement and educational attainment using large nationally representative longitudinal studies of students (the National Educational Longitudinal Study, NELS:88, and the National Longitudinal Survey of Youth, NLS:79), which began in 1979. Hooyse compared the percent of the variation in student achievement on various field tests (math, reading) explained by family, school, and community factors. For every test, the percent of the variation explained by the family variables for elementary school input variables for 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.

Moreover, this is not a problem that is "fixed" by meta-analyzing large numbers of flawed cross-section studies, since 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 bullseye (the true effect). Firing more shots will simply do a better job of identifying the point around which the sights is targeted but will not help determine where the bullseye is. Thus requires that the bias or error in the rifle scope be fixed.

5. 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 considers cross-section as well as descriptive studies. Nonetheless, he finds a best topical support research for aggressive regulation of the teacher labor market. On the question of whether pedagogical training contributes to teacher effectiveness, he finds only a "limited" support in the research, and adds: "It is clear that the evidence reviewed for this report, however, whether such knowledge and skills are best acquired through coursework, field experience (especially student teaching) or on the job" (Allen, 2001, p. 29). On the question of whether more stringent screening for teacher training program entrants pays off in terms of student achieve-