

Improving Academic Performance in U.S. Public Schools:  
Why Teacher Licensing is (Almost) Irrelevant

Michael Podgursky  
Department of Economics  
University of Missouri - Columbia

Prepared for “Teacher Preparation and Quality: New Directions in Policy and Research”  
Conference, American Enterprise Institute. Washington D.C., October 18-20, 2003

## Introduction: Six Propositions About Teacher Labor Markets

Participants in this conference have been asked to assess several possible reforms for teacher licensing. Two ways of “raising the bar” are being considered. The first focuses on more rigorous testing of teachers. The second considers more rigorous standards for schools of education. These reform models are summarized in Figure 1. Moving down the rows indicates more rigorous screening based on tests. Moving along the columns indicates greater screening of teacher training institutions, presumably by state education agencies, accrediting organizations such as NCATE or both. The southeast cell in the matrix would have more of both, and perhaps pick up interactions between the two reforms, e.g., higher testing bars might enhance the return from higher standards for teacher program approval. All three of these approaches amount to a restricting of supply into the teaching labor market as compared to the current system.

My charge is to evaluate the “low-low” cell in the Figure 1 matrix, i.e., relaxed state regulation of teacher training programs and low barriers for individuals. A thumbnail sketch of the “low-low” (or, as I would prefer, “flexible”) regime follows. State education agencies would 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 they specialize, but the bar would not be particularly high. In this regime, then, the focus of state regulation would be on student learning and not teacher training and licensing. Parents would be protected from incompetent teacher practice through state monitoring of student learning and though school choice. Private associations 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.

All of the authors have been asked to address a series of questions concerning their assigned regime. In order to lay the groundwork for a discussion of the flexible regime, I begin with some observations concerning research 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.

There is widespread consensus in the research community that scientific evaluation of social policy programs (including education) requires: a) randomized experimental study design, or b) non-experimental longitudinal data on participants.<sup>1</sup> Unfortunately, little research on teacher testing or licensing meets either standard and the research that does is tentative and inconclusive.

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. Unfortunately, at present there is no research on teacher credentials or training that meets this standard, although the Institute for Education Sciences of U.S. Department of Education is promoting such studies (Moesteller and Boruch, 2002; U.S. Department of Education, undated) and some are under way. Thanks to these efforts, five 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 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 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 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 (Sanders and Horn, 1994; Hanushek, 1992; Hanushek, Kain, and Rivkin, 1999; Rivkin, Hanushek, and Kain, 2000; Aronson, Barnow, and Sanders, 2003).

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 background of students in classrooms and student SES is correlated with teacher credentials. In the language of econometrics, we would say that these cross-section studies of teacher credentials suffer from "omitted variable bias."

A recent study by Hoxby (2001) highlights the importance of these socioeconomic variables and their potential for producing bias in teacher effects research. Hoxby analyzed the effect of family, neighborhood, and school input variables on student achievement and educational attainment using two large nationally representative longitudinal studies of students (the National Educational Longitudinal Survey, NELS88, and the National Longitudinal Survey of Youth, which began in 1979). The list of variables included in each of the areas is extensive. Family variables include parent's education, family income, student race and ethnicity, books at home, etc. The school input variables include per-pupil spending, average class size, average teacher salary, maximum teacher salary, percent of teachers with MA's, average experience of the teacher, teacher certification status, and other information on school resources. Community variables include income and demographic data on households in the school district and city.

Hoxby compared the percent of the variation in student achievement on various field tests (math, reading) explained by each of these sets of factors. For every test, the percent 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.

Consider teacher licensing. Given the complexities of teacher licensing systems, virtually every school district in the U.S. has some teachers out of compliance, however, substandard certification tends to be relatively more common in schools with low SES students. Since 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. 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.<sup>2</sup>

The number of studies of teacher certification that meet these 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 SES. The results of these studies (both by Goldhaber and Brewer 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 (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 percent found no statistically significant effect. Of the three finding a significant effect, two were positive and one was negative.

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 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 finds only “limited” support in the research, and adds: “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.” (Allen, p. 29). On the question of whether more stringent screening for teacher training program entrants pays off in terms of student achievement, he finds the literature “inconclusive.”

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 one gauges the contribution of teachers in the upper and lower tails of the teacher effects distribution to student achievement gains the effect is often quite substantial. However, these teacher effects are largely unrelated to traditional measures of teacher quality such as test scores, certification credentials, experience, or graduate degrees, a result highlighted in a recent survey by Goldhaber (2002). Hanushek and Rivkin (2003), summarizing their own and other research come to the same conclusion. A new study of Chicago public teachers by Aaronson, Barrow, and Sander (2003) uses a very large longitudinal file of linked student achievement scores. What makes this study unique is that the authors also have very extensive administrative data on teacher characteristics that are unavailable in other studies, including education, experience, types of teaching licenses, and selectivity of the teacher's undergraduate college. They find that over ninety percent 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.

### 3. 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, while the evidence for any particular intervention is weak, why not implement a variety of such measures, on the chance that one or two 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?

As I have indicated above, the evidence for benefits from "raising the bar" is very weak.

However, if these reforms were costless, then one might make the case for their implementation on the chance that some benefits would accrue. However, raising entry barriers produces costs as well as (potential) benefits – a point often overlooked in education policy debates.

First there are the direct resource costs. To the extent that we raise requirements for education coursework, we incur direct educational costs. Annual costs per student in higher education currently are roughly \$27,000 dollars per year. 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 a year and a half of teaching courses (including student teaching), this is a very costly investment. Even at the minimum wage, this amounts to over fifteen thousand dollars. One perverse result of requiring seat time in pedagogy courses for labor market entry is that candidates with greater academic skills, who presumably have a higher alternative wage, face a higher costs in securing a teaching license (Ballou and Podgursky, 1997).

Licensing also imposes other less obvious costs. Licensing means that school districts cannot hire an unlicensed candidate as long as a licensed candidate is available. Some advocates would prevent school districts from ever hiring an unlicensed (or under-licensed) candidate. Darling-Hammond (2003), for example, argues that the state has a constitutional duty to provide a fully-licensed teacher in every public school classroom in California. Thus, one important cost of an increase in a licensing entry barrier is the reduction in the size of the applicant pool from which a school administrator may hire as a result of the restriction.

School administrators know many things about teaching 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.<sup>3</sup> By preventing school administrators from considering any unlicensed applicants, school districts are forced to hire the worst certified candidate even if a superior non-certified 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 pool, so they have less ability to select among candidates on the basis of the factors that they observe but the regulators do not.

This point is illustrated in Figure 2. Here I have presented hypothetical data on the distribution of teacher quality among certified and uncertified applicants. While these data are hypothetical, I believe that they represent the general picture that is emerging in the “teacher effects” literature, namely, that the dispersion of teacher effects is large relative to any measured teacher characteristic such as teacher 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 certification effect of sufficient size that the average certified teacher is better than 60 percent of non-certified teachers. I see this as an upper end estimate for certification. However, the conclusions which follow do not hinge in any significant way on this assumption. If we bumped the certification premium mean up to 70 percent none of our basic conclusions change. The key point is that there is a large dispersion of quality within the certified and uncertified pools – a stylized fact strongly supported in the research literature.

(Figure 2)

Suppose Figure 2 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 is drawn at random from the uncertified and the certified pool. What is the probability that the certified candidate is the superior teacher? In Table 1 we see 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 67<sup>th</sup> percentile of the certified distribution. Now suppose the school has five random applicants from the certified population, but no uncertified applicants. Obviously, the probability of hiring a certified candidate is 100 percent. However, note that the quality of the best applicant jumps sharply, from the 67<sup>th</sup> to the 88<sup>th</sup> percentile. This illustrates an important point. If teachers are screened well (a point taken up below), a larger applicant pool means higher quality hires. 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 two hundred applicants as compared to twenty.

(Table 1)

Can this district benefit from access to a pool of lower average quality uncertified applicants? The statistics in the bottom row of the table show that the answer is yes. Here we suppose that five uncertified applicants are allowed to apply for the job along with five certified applicants. Even though the uncertified applicants are of lower average quality, 39 percent of time the uncertified applicant will be the best of the ten applicants. The last column shows that by expanding the applicant pool from five to ten 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. Now we go from row three with ten applicants to row two with five

applicants. Average teacher quality falls, in spite of the fact that the average certified teacher is better than the average non-certified teacher. The fall in mean teacher quality as we go from row three to row two illustrates the hidden cost of a licensing entry barrier: shrinking the applicant pool reduces the mean quality of the resulting hires.

As noted above, this general point would not change if we raised the relative mean quality of certified teachers. What produces this benefit is the fact that sometimes a non-certified teacher is the best candidate. A tight licensing entry barrier prevents school districts from ever hiring unlicensed teacher.

A 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 (2000) assumes that school administrators have imperfect, but independent, information about the quality of job applicants (i.e., over and above certification or a test score). A similar result obtains. Ballou also shows that the cost of the reduced supply tends to be larger for low SES districts which tend to draw relatively more applicants around the cut scores.

To summarize, “raising the bar” on teacher licensing shrinks the available applicant pool available to schools. Other things equal, this will tend to lower 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. Thus, we are faced with the very real risk that mean teacher quality might actually decline as a result of such policies.<sup>4</sup>

#### 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 often confronted with the question: “would you send your children to an unlicensed doctor?” Although rarely stated explicitly, I believe the argument is as follows: “While it may be true that there exists no rigorous evidence for the reforms we have proposed (e.g., more rigorous teacher testing, accreditation and more rigorous review of teacher training institutions), they are broadly similar to what is found in medicine. Therefore if we implement them, 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 why I think that the 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 argued that some sort of 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, 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 the humanities than to medicine or the sciences. Controlled experiments and randomized studies are rare. Use of large-scale longitudinal data on students is not widespread. However, what faculty at leading research universities do or do not do is largely irrelevant since 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 practitioners and their professional associations are in no position to absorb this research and incorporate it into their teaching or their standards. Teacher professional 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, i.e., 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 (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 and consumers in the market for education services seem quite capable of making informed choices among practitioners. While 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.<sup>5</sup> 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 pervasive evidence 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 where unlicensed practice is the norm. The latter markets seem to work quite well with little or no government regulation. Indeed, my own observation of these markets over many years as a labor economist leads me to conclude 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 and 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, etc., the armed services have taught these young recruits basic literacy and numeracy skills as well. 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, and covers material that students should have learned in elementary and secondary schools. States do not require that the faculty in community colleges be licensed, and evidence suggest that most are not certified teachers. Nonetheless, if we judge success by enrollment growth, these institutions 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 U.S. that long predates the public school system. Private schools routinely hire unlicensed teachers. Figure 3 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 non-religious schools, where just 55.9 percent of teachers are certified. The rates are lower still at the secondary level. In non-religious secondary schools the certification rate is just 35.1

percent. Thus, while private schools do hire certified teachers, they also hire substantial numbers of non-certified teachers as well. It should also be noted that charter schools, too, hire large numbers of uncertified teachers.

(Figure 3)

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 the student academic achievement (Winkler, 1975; Summers and Wolfe, 1977; Ehrenberg and Brewer, 1993, 1994). 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 almost consistently higher in the non-certified population. A similar pattern is seen in charter schools.

(Table 2)

#### 6. State Teacher Licensing Systems Are So Complex That No One Is in Compliance Anyway

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 non-vocational). 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.

How is it that the public is protected by a single license in other professions, yet K-12 education requires over one hundred? 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 one 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. (When there is an accident in a public place, one hears “Is there a doctor in the house?,” not a call for a medical specialist.) 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.” 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.

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, 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 of the teacher of record.

The excessively complex licensing system in K-12 education can also be seen as a means of rent capture by teacher unions and schools of education. 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 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 a compromise is to proliferate certificate areas. This generates more demand for education school courses and but also restricts supply to school districts.

As a consequence of the complex licensing systems that states have constructed, virtually no school district in the U.S. is in full compliance. The complexity of the state licensing systems make national tabulations of unlicensed, uncertified, or sub-standard certification difficult. Thus I will illustrate this point with administrative data from two states. Figure 4 presents data for Missouri public K-12 school districts (I have excluded k-8 districts). On the vertical axis we have plotted the percent of courses taught by teachers with inappropriate licenses during the 2001-2002 school year. On the horizontal axis we have plotted spending per student in average daily attendance. First note that of 447 K-12 school districts not a single district had no courses taught by an inappropriately licensed teacher. Second, prevalence of inappropriate licensed practice would seem to have little to do with school spending. In fact, the correlation between the rate of unlicensed teaching and spending per student is positive and statistically significant (.27).

(Figure 4)

My second example is decidedly non-random. Westchester County New York is home to some of the wealthiest households and highest paid school teachers in the U.S. The schools in these exclusive communities might appropriately be compared to the very best private schools. In Figure 5 I have plotted percent uncertified teachers and median teacher salaries for school year 2000-2001 from the most recent report of the NYS Department of Education. (Uncertified means the percent of classroom teachers who either teach more than 20 percent of their time in a

subject or subjects for which they hold no certification or who hold an emergency license.) In spite of very high salaries, no district in Westchester County has fewer than two percent of their teachers uncertified. Note that Scarsdale, which boasted a 2000-2001 median teacher salary of \$90,191, had six percent of its teacher 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 U.S. is in full compliance with the New York State law this raises serious questions as to how we can seriously contemplate “raising the bar” and further restricting supply.

(Figure 5)

## II. The Model

With the above as background, the case for the “flexible” model is relatively straightforward. State education regulators protect the public by focusing on what they can measure (student learning), and not on what they cannot (teacher quality). That is, they focus on education outputs rather than education inputs. As noted above, 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. Thus, the model that emerges has several features.<sup>6</sup>

The first, and 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 performance in terms of student learning increasingly face administrative sanctions. Another important mechanism for producing accountability is school choice. Indeed, one of the best 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, the role of the state in monitoring “teacher quality” – something that they can’t measure anyway – should wither 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. Teachers should also be required to hold a bachelors 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 (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 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 and make the implausible argument that the scores only have “validity” for licensing but not hiring. Imagine if colleges or

professional schools received similar information. Continuous SAT or GRE scores would be collapsed into binary “good enough for graduate study” and “not good enough for graduate study” indicators. 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.

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 renew teacher contracts during an initial probationary period of two to five 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.

Even if a “magic bullet” in terms of teacher training or testing were found it, would be many years before new, more effective teachers would have diffused through the teaching workforce. Teacher turnover averages roughly six percent annually and about half of new teacher hires are returning teachers or interdistrict transfers. Thus, only a very small share of teachers are affected by licensing reform in any year. Schools have information on the job performance of the 90 percent 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 anti-competitive practices on the part of teacher unions, schools of education, or other education producer organizations.

- Minimal state licensing standards for teachers: criminal background check, bachelors 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 the 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 program in California are examples. In such programs, prospective teachers must hold a BA, pass exams and demonstrate content knowledge to receive a probationary license. After a probationary period of two to three years (which in these states includes on or off-site professional development), mentored practice, and satisfactory evaluations by supervisors, the teachers receive a standard license. Such alternate route teachers satisfy the “highly qualified teacher” requirement of the No Child Left Behind Act.

### Role of Private Organizations

Some years ago, Myron Lieberman perceptively described the k-12 education industry as “producer-dominated” (Lieberman, 1996). 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 above 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 wither 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 Operations Research or Computer Science majors from the engineering school. There are no licensing entry barriers creating a monopoly for a particular college. Opening up the teacher training market to competition is a much 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 truly mediocre schools of education have succeeded in securing approval from state education agencies as well as accreditation from NCATE.

Finally, this will require a 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 the change of gestalt is required. That 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. 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 or not 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 spent to lower class sizes, or otherwise enhance student welfare have been wasted. 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 pre-school care.

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 “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 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, relaxing licensing standards will 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 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 a much more natural variation in the workforce in the credentials and training of teachers. This will permit better evaluation of the effects of credentials.

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 such 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 non-certified 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, presumably there would be a much larger pool of the former. With a larger pool from which to choose, those hired would presumably be of higher quality. As noted above, on average schools will end up with a better hire if they have fifty applicants for a job than if they have five.

### III. Conclusion: Teacher Licensing Needs to Whither Away

Policy debates about “teacher quality” have unfortunately 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 we were to reform teacher licensing or training so as to raise student achievement, 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, in other professional labor markets the quality of performance for incumbents is primarily determined by incentives: experienced dentists who do

a poor job on our teeth lose customers; experienced dentists who perform incompetently lose get sued.

Rather than dwell on the credentials and training of the 3-4 percent newly minted teachers hired 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 make use of available information on teaching performance for the 95 percent of incumbent teachers. Dismissing 3-4 percent 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 3-4 percent of newly minted hires.

However well intentioned, attempts to address the teacher quality problem by “raising bars” in teacher licensing are likely to make things worse rather than better. Since 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 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 than fewer.

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

## References

Allen, Michael. B. 2003. Eight Questions on Teacher Pay: What Does the Research Say? Denver, CO: Education Commission of the States.

Angrist, Joshua, and Jonathan Guryan. 2003. "Does Teacher Testing Raise Teacher Quality? Evidence from State Certification Measurements." Cambridge, MA: NBER Working Paper 9545 (March).

Armor, David, Patricia Conry-Osenguera, Millicent Cox, Nicelma King, Lorraine McDonnell, Anthony Pascal, Edward Pauly, and Gail Zellma. 1976. Analysis of the School Preferred Reading Program in Selected Los Angeles Minority Schools. Santa Monica, CA: RAND .

Aaronson, Daniel, Lisa Barrow, and William Sander. 2003. "Teachers and Student Achievement in the Chicago Public High Schools" Working Paper. Research Department. Federal Reserve Bank of Chicago.

Ashton, Patricia, and Linda Crocker. 1987. "Systematic Study of Planned Variations: The Essential Focus of Teacher Education Reform." Journal of Teacher Education. (May-June), pp. 2-8.

Ballou, Dale. 1999. "Recruiting Smarter Teachers: Is Testing the Answer?" Vanderbilt University.

Ballou, Dale and Michael Podgursky. 1997. Teacher Pay and Teacher Quality. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Ballou, Dale and Michael Podgursky. 1998. "Teacher Recruitment and Retention in Public and Private Schools." Journal of Policy Analysis and Management. Vol. 17 No. 3 (Summer), pp. 393-418.

Broughman, Stephen P. and Mary R. Rollefson. 2000. Teacher Supply in the United States: Sources of Newly-Hired Teachers in Public and Private Schools, 1987-88 to 1993-94. Washington, DC: National Center for Education Statistics 2000-309.

Darling-Hammond, Linda. 2003. "Access to Quality Teaching: An Analysis of Inequality in California Public Schools." [http://www.mofo.com/decent\\_schools/expert\\_reports/darling-hammond\\_report.pdf](http://www.mofo.com/decent_schools/expert_reports/darling-hammond_report.pdf)

Ehrenberg, Ronald C. and Dominic J. Brewer. 1993 "Did Teachers' Race and Verbal Ability Matter in the 1960's? Coleman Revised" Economics of Education Review Vol 14 No. 1, pp. 1-23.

Ehrenberg, Ronald C. and Dominic J. Brewer. 1994. "Do School and Teacher Characteristics Matter? Evidence from High School and Beyond." Economics of Education Review. Vol. 13. No. 1, pp. 1-17.

Ehrenberg, Ronald C. and Robert S. Smith. 1996. Modern Labor Economics. 6<sup>th</sup> Ed. Reading MA: Addison-Wesley.

Evertson, Carolyn M., Willis D. Hawley, Marilyn Zlotnik. 1985. "Making a Difference in Educational Quality Through Teacher Education." Journal of Teacher Education (May-June), pp. 2- 12.

Ferguson, Ronald F. 1991. "Paying for Public Education: New Evidence on How and Why Money Matters." Harvard Journal on Legislation Vol 28, pp. 465-498.

Ferguson, Ronald F. and Helen Ladd. 1996. "How and Why Money Matters: An Analysis of Alabama Schools." in Helen Ladd (ed.) Holding Schools Accountable: Performance-Based Reform in Education. Washington, DC: Brookings Institution.

Feistritzer, Emily C and David C. Chester. Alternative Teacher Certification: A State-by-State Analysis: 1998-99. Washington DC: National Center for Education Information.

Goldhaber, Dan. D. and Dominic J. Brewer. 1997. "Why Don't Schools and Teacher's Seem to Matter?" Journal of Human Resources Vol. 32 No. 3 (Summer), pp. 505-523.

Goldhaber, Dan. 2002. "The Mystery of Good Teaching." Education Next Vol. 2 (Spring). pp. 50-55.

Greenwald, Rob, Larry V. Hedges, and Richard D. Laine. 1996. "The Effect of School Resources on Student Achievement." Review of Education Research. Vol. 66, No. 3 (Fall), pp. 361-96.

Hanushek, Eric A. 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." Journal of Economic Literature. Vol. 24. No. 3 (Summer), pp. 557-77.

Hanushek, Eric. A. 2003. "The Failure of Input-Based Resource Policies." The Economic Journal. Vol. 113 No. 485 (February). pp. F64-F98.

Hanushek, Erik A. and Steven G. Rivkin. 2004. "How to Improve the Supply of High Quality Teachers" Brookings Papers in Education Policy: 2004. Washington, DC: Brookings Institution. Forthcoming.

Heckman, James J., Robert J LaLonde, and Jeffrey A. Smith. 1999. "The Economics and Econometrics of Active Labor Market Programs." In Orley Ashenfelter and David Card (eds.) Handbook of Labor Economics. Netherlands: Elsevier Science, pp. 1865-2097.

Hess, Frederick M. 2001. Tear Down This Wall: The Case for Radical Overhaul of Teacher Certification. Washington DC: Progressive Policy Institute (November).

Hoxby, Caroline. 2001. "If Families Matter Most, Where Do Schools Come In?" in Terry M. Moe (ed.) A Primer on America's Schools. Stanford University: Hoover Institute Press.

Jacob, Brian A. and Lars Lefgren. 2002. "The Impact of Teacher Training on Student Achievement: Quasi-Experimental Evidence from School Reform Efforts in Chicago." NBER Working Paper 8916. Cambridge, MA: National Bureau of Economic Research.

Lieberman, Myron. 1995. Public Education: An Autopsy. Cambridge, MA: Harvard University Press.

Mosteller, Frederick and Robert Boruch (eds.) 2002. Randomized Trials in Education Research. Washington DC: Brookings Institution.

Murnane, Richard. 1975. The Impact of School Resources on the Learning of Inner City Children. Cambridge, MA: Ballinger.

National Commission on Teaching and America's Future. 1996. What Matters Most. Columbia University, Teachers College.

\_\_\_\_\_. 1997. Doing What Matters Most. Columbia University, Teachers College.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 1998. "Teachers, Schools, and Academic Achievement." Unpublished Manuscript.

Rottenberg, Simon. 1962. "The Economics of Occupational Licensing." in National Bureau of Economic Research. Aspects of Labor Economics. Princeton, NJ: Princeton University Press.

Sanders, William L. and Sandra P. Horn. "The Tennessee Value-Added Assessment System (TVAAS): Mixed Model Methodology in Educational Assessment." Journal of Personnel Evaluation in Education. Vol. 8, pp. 299-311.

Summers, Anita M. and Barbara L. Wolfe. 1977. "Do Schools Make a Difference?" American Economic Review Vol. 67, No. 4 (September), pp. 639-652.

U.S. Department of Education. Institute of Education Sciences. Undated. Random Assignment in Program Evaluation and Intervention Research: Questions and Answers  
<http://www.mathematica-mpr.com/PDFs/randomassign.pdf>

Winkler, Donald R. 1975. "Educational Achievement and School Peer Group Composition." Journal of Human Resources. Vol. 10, No. 3 (Spring), pp. 189-204.

Table 1

Probability That a Non-Certified Candidate is the Best Applicant

Applicant Pool	Probability that Certified Applicant Is Best Candidate In Pool	Average Quality Of Best Applicant, %-ile of Cert. distr.
1 Certified 1 Uncertified	.57	67
5 Certified 0 Uncertified	1.00	88
5 Certified 5 Uncertified	.61	92

Table 2

Measures of Teacher Quality in Public and Private Schools

	Traditional Public	Public Charter		Private Religious		Private Non-Religious	
	All	Certified	Not Certified	Certified	Not Certified	Certified	Not Certified
College Selectivity: Most Competitive	.9	1.4	4.0	.8	3.1	3.4	9.4
Other Selective	22.1	26.7	29.5	18.5	21.4	30.3	32.8
Total Selective	23.0	28.1	33.5	19.3	24.5	33.7	42.2
Math and Science Majors	10.1	10.1	12.8	11.1	10.8	11.6	13.7
Academic Majors	33.7	40.9	56.3	37.2	55.5	39.0	63.0

Source: 1999-00 Schools and Staffing Surveys

Figure 1

Regulatory Regimes

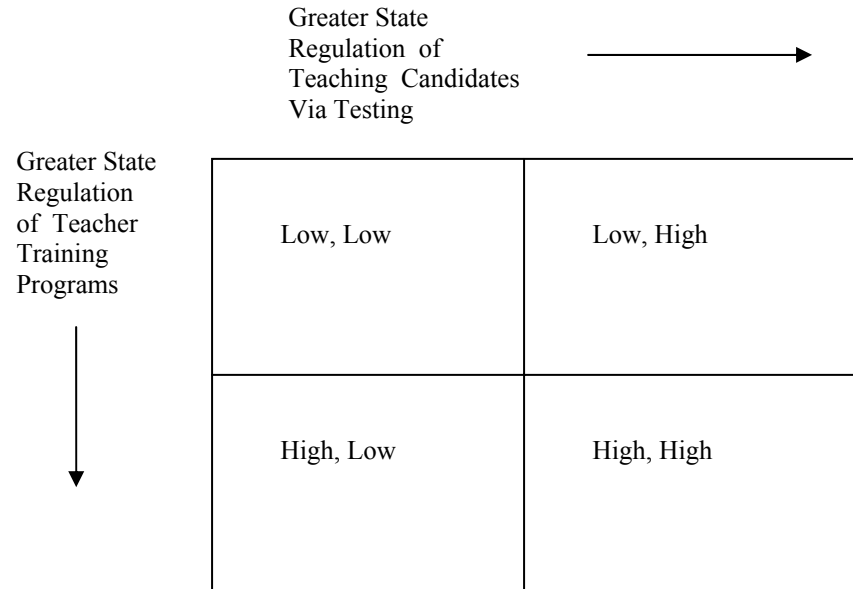


Figure 2: Overlapping Ability When The Average Certified Applicant Is Better Than 60 Percent of Non-Certified Applicants

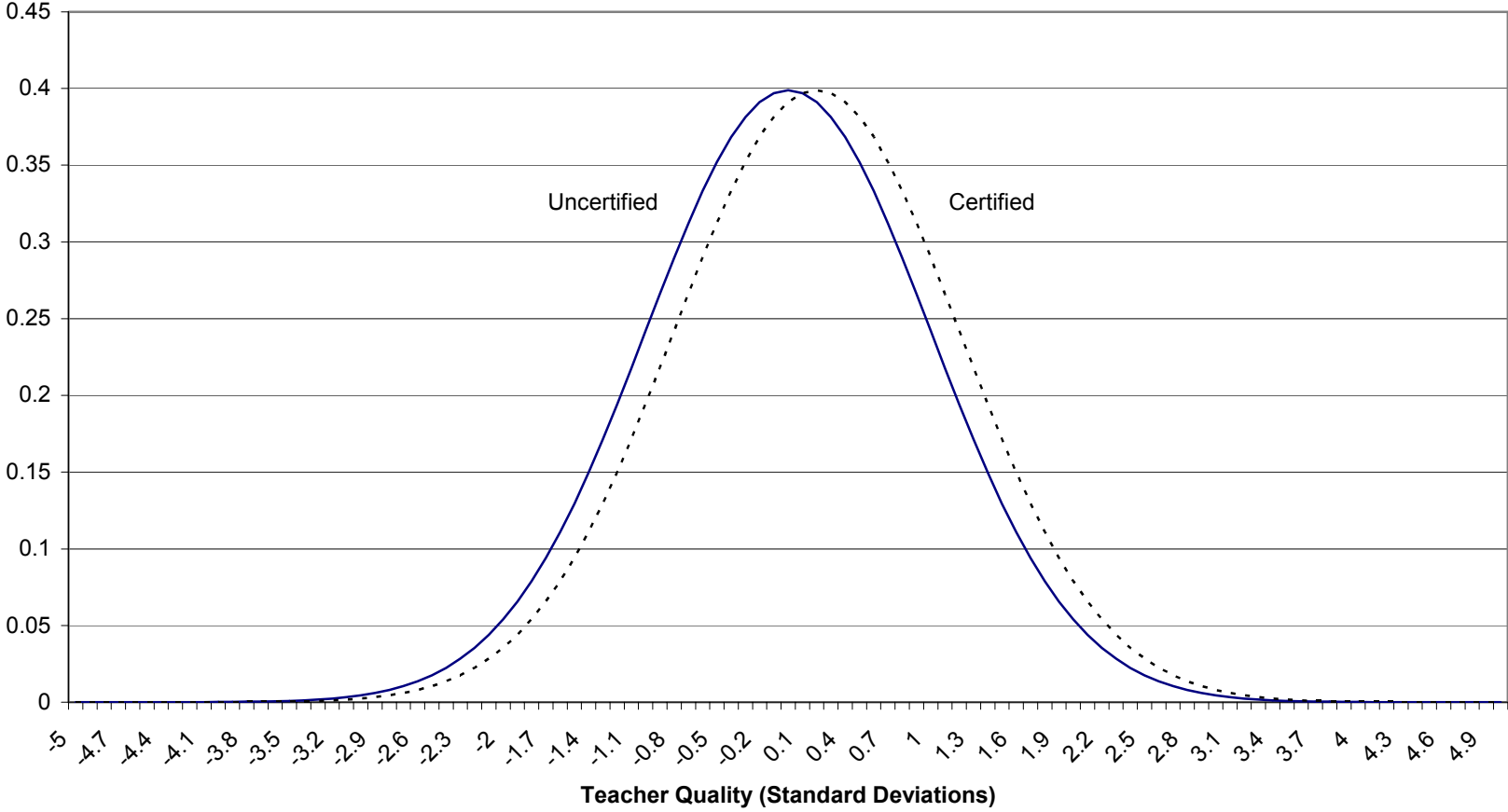
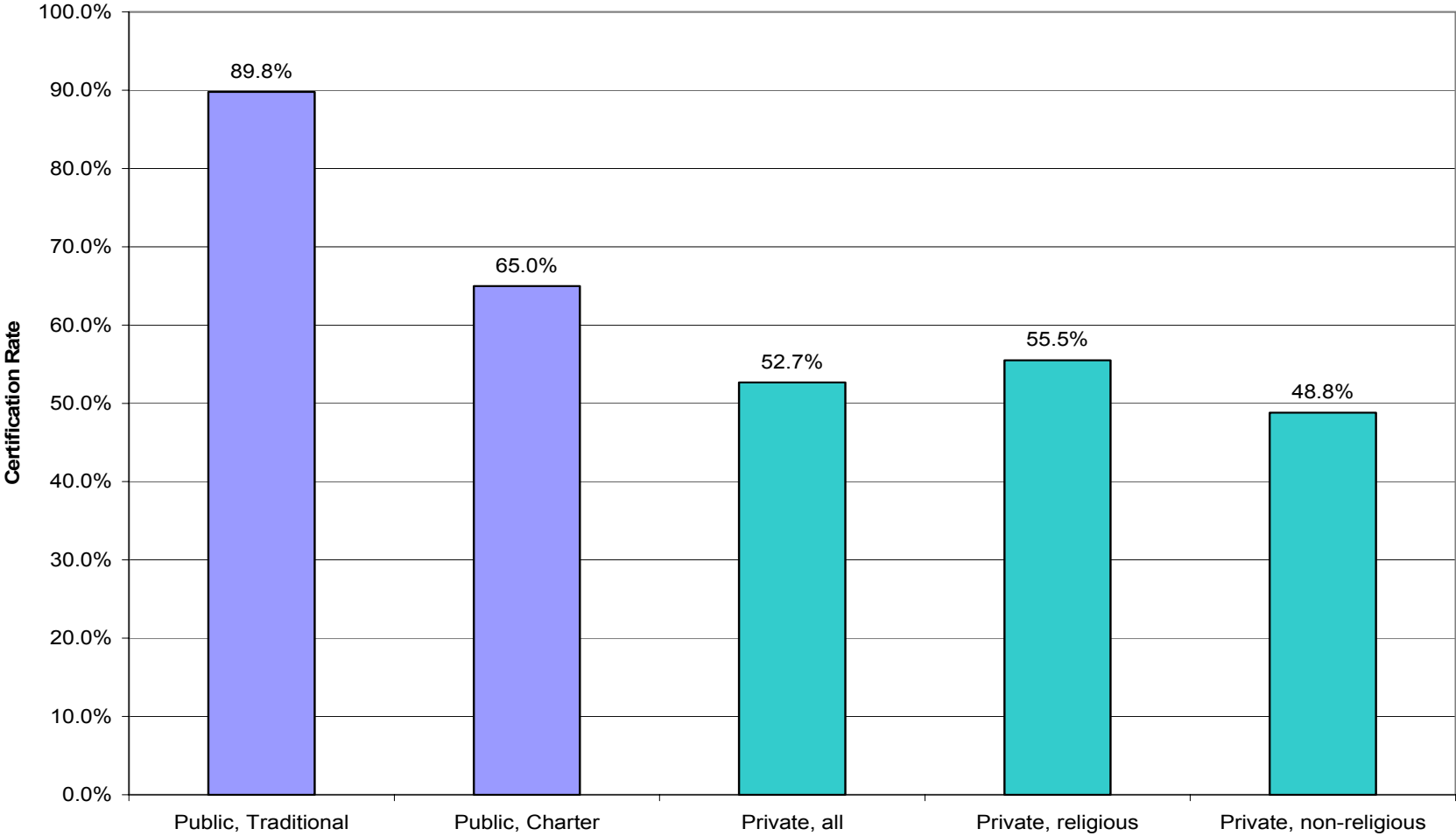


Figure 3: Percent of Teachers Holding Regular State Certification in Primary Teaching Area: Traditional Public, Private, and Charter Schools



Source: 1999-00 Schools and Staffing Surveys

Figure 4

Percent of Courses Taught by Teachers With Inappropriate or No Licenses by Expenditure Per Pupil in Average Daily Attendance:  
Missouri K-12 Public School Districts, 2001-2002

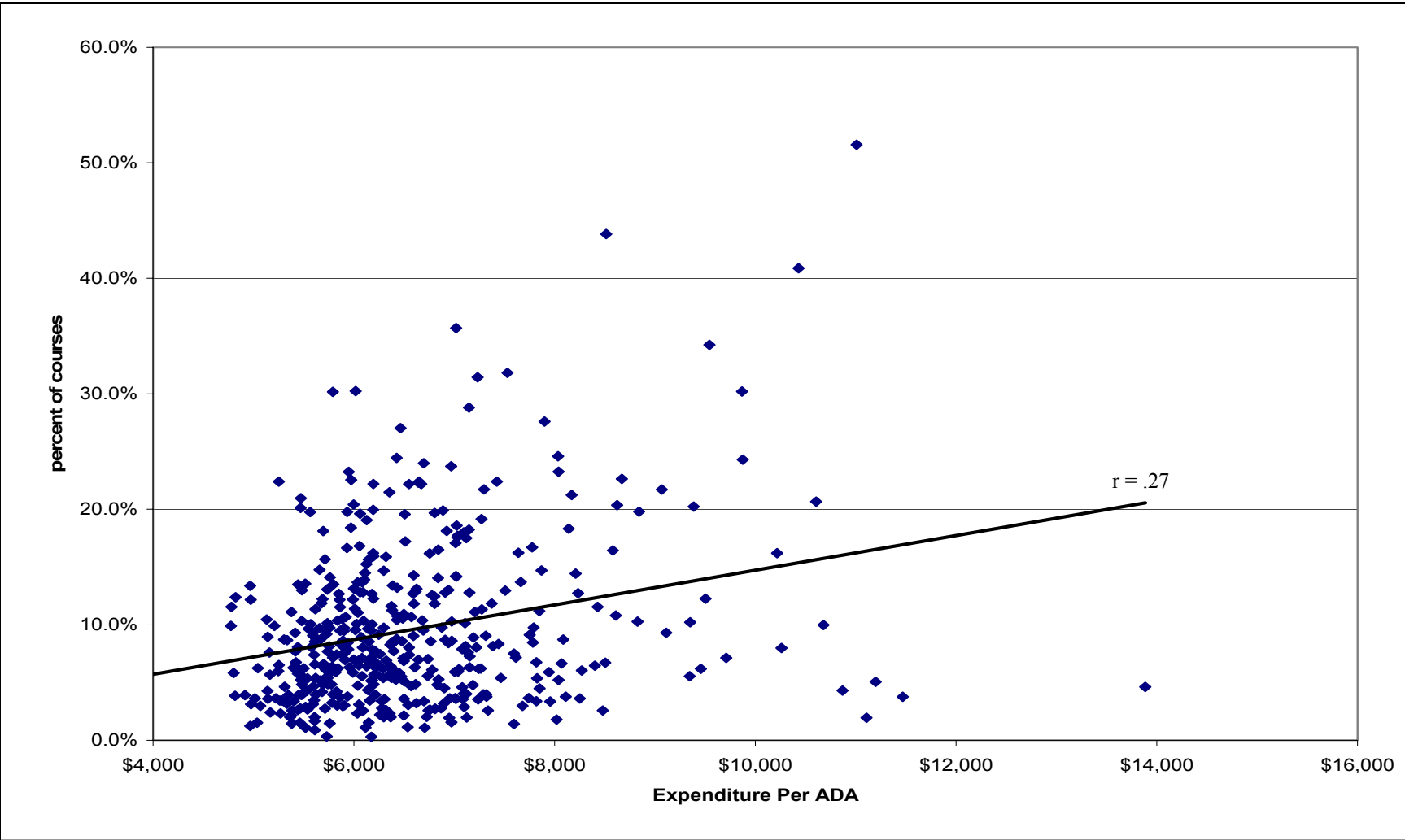
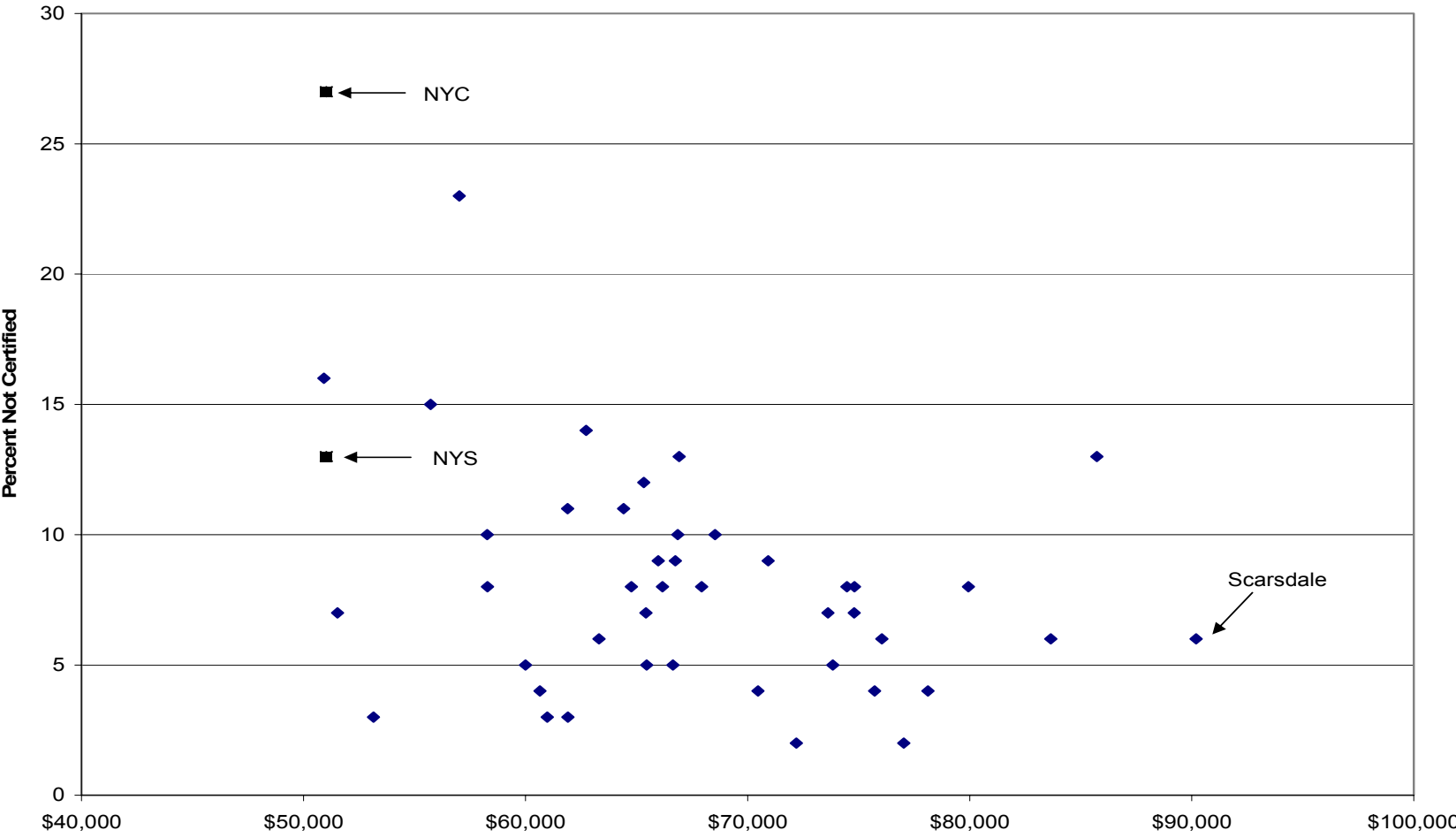


Figure 5

Percent of Teachers Not Certified and Median Salaries in Westchester County, NY Public School Districts: 2000-2001



## References

---

<sup>1</sup> 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 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 “quasi-experiment” to examine the effect of teacher training on student achievement. However, even with “natural experiments” longitudinal data is highly desirable.

<sup>2</sup> 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.

<sup>3</sup> Studies using student longitudinal data by Armor, et.al. 1976, Murnane, 1975 find large effects of principal evaluations on student achievement gains. More recently Sanders and Horn (1994, p. 2000) report : “There is a very strong correlation between teacher effects as determined by the data and subjective evaluations by supervisors.”

<sup>4</sup> Ballou (2000) shows 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.

<sup>5</sup> A recent paper by Angrist and Guryan (2003) find 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 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 profession 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 NCTM 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 regulation coming from private school consumers.

<sup>6</sup> The model proposed here is similar that in Hess (2001), which he describes as “competitive certification.”