

The Two-Legged Stool: The Neglected Role of Educational Standards in Improving America's Public Schools

Julian R. Betts

Two of the most important reforms to American public schooling in this century have been an increase in the minimum school-leaving age and a dramatic increase in expenditures per pupil. The former reform has generally been hailed as a success, given evidence that an extra year of schooling significantly boosts students' earnings later in life. However, evidence on the effectiveness of the trend toward higher spending per pupil, smaller class sizes, and more highly educated and trained teachers is much more mixed. A host of studies on the link between school finances and test scores has not shown a systematic link between spending and achievement. Another set of studies tests whether higher school spending leads to higher earnings for students later in life. The findings in this body of work are also mixed: even the most optimistic

results suggest a very low rate of return to increased school expenditures.

Given the central role of public schools in preparing younger generations for the workforce and the sizable expenditures devoted to public schools, it becomes important to ask whether other reforms might increase the effectiveness of public schooling. The premise of this paper is that educational standards are a key element in school reform. The paper suggests that the twin policies of higher school-leaving ages and higher spending would have been much more effective if accompanied by systematic increases in educational standards. In a sense, these two policies form a two-legged stool. Increasing the number of years that students must spend in school and increasing spending per pupil—without at the same time stipulating what subject matter students should have mastered by each age—are unlikely to maximize the rate at which students learn. The missing “leg” in these past reforms is a set of academic standards against which both students *and*

Julian R. Betts is a professor of economics at the University of California, San Diego.

schools are measured. Instead of allowing students to drop out of school merely because they have reached the age of sixteen, why not require students who wish to leave school before graduation to pass a set of minimum competency tests? Similarly, does it make sense to increase funding for public schools without at the same time requiring the schools to document that they have translated these additional expenditures into improved student achievement? In both cases, the need for a clear set of curriculum standards, backed by testing, is clear.

The central point is simple: far and away the most important determinant of how quickly students learn is the effort of students themselves. It follows that an increase in schools' expectations of students could have important effects on the quality of public schooling. By establishing a rigorous set of educational standards, schools can create a set of incentives and rewards to promote student learning. Higher standards can increase the effectiveness of school-leaving policies. Instead of simply allowing students to drop out at age sixteen, schools might require an exit exam. In this way, a minimal level of academic achievement, rather than age alone, would become the prerequisite for dropping out of school. Similarly, reductions in class size might become more effective if, at the same time, schools increased the standards that students at each grade level are expected to meet.

The next section briefly outlines the history of past reforms related to the minimum school-leaving age and spending per pupil. The subsequent section outlines ways in which schools can heighten their expectations of students. In practice, "higher expectations" can come in the form of additional homework, the development of curriculum standards in conjunction with an assessment of students' progress in mastering the curriculum, stricter grade-promotion policies, and more stringent grading of students. Later sections then analyze whether such policies to promote higher educational standards work in practice. The penultimate section delves into practical and political issues that can afflict a school administration when it tries to increase standards, and suggests solutions to some of these problems.

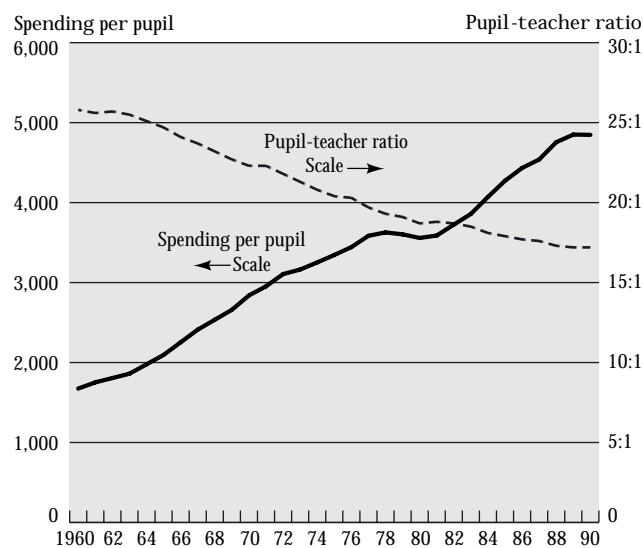
A REVIEW OF THE IMPACT OF PAST INCREASES IN THE SCHOOL-LEAVING AGE AND IN SPENDING PER PUPIL

Between 1960 and 1990, real spending per pupil and the teacher-pupil ratio increased dramatically (Chart 1). (Expenditures per pupil are expressed in 1990-91 prices, and so account for inflation.) Betts (1995a), among others, documents that there has been a strong trend toward increased spending per pupil throughout the century in the United States. This trend represents one of the most important changes in the recent history of public schooling.

A second important reform in American public schooling has been increases in the school-leaving age. Lang and Kropp (1986) document that over time the number of states without a school-leaving age has dropped, and the average school-leaving age has risen considerably (Chart 2). Most of the changes in the school-leaving age occurred in the first half of the century. (However, Lang and Kropp note that by 1965 two states—Mississippi and South Carolina—had abolished their compulsory attendance laws.)

Chart 1

Trends in Real Spending per Pupil and in the Pupil-Teacher Ratio



Source: National Center for Education Statistics (1991).

Note: Spending per pupil is in 1990-91 prices.

What benefits, then, have resulted from these two important transformations in American public schooling? Consider first the impact of additional school spending. For over three decades, social scientists have examined the link between school expenditures and student success. Most of this effort has modeled test scores as a function of spending per pupil, the teacher-pupil ratio, and other measures of school inputs. Since other contributors to this conference volume address this literature, I will not review it in detail here. It suffices to mention that in a recent review, Hanushek (1996) found that of 163 estimates of how spending per pupil affects student performance, only 27 percent found a positive and significant relationship. Similarly, of 277 reported estimates of the impact of the teacher-pupil ratio on student performance, only 15 percent found a positive and significant link, while 13 percent reported a negative and significant link. These figures do not mean that money never matters. Instead, they suggest that the relationship, if positive, is a rather tenuous one. In American schools, at least as they have operated in the past, spending has not had large or systematic effects on student achievement.

The conclusion drawn from the statistical research is supported by aggregate trends in school spending and in student achievement. Chart 1 shows trends in current expenditures per pupil and the pupil-teacher ratio in American public schools between 1960 and 1990. By both measures, the financial resources spent on public school students have risen markedly over the last three decades. Yet during the same period, student achievement has hardly changed, and by one measure it may even have fallen. Test scores on the National Assessment of Educational Progress, a test given to a random sample of students in various grades since the early 1970s, have changed little over the 1970s and 1980s. Trends in the Scholastic Aptitude Test show a sharp decline in the late 1960s, a more gradual decline during the 1970s, and a partial recovery since then (see, for instance, Hanushek [1996]).

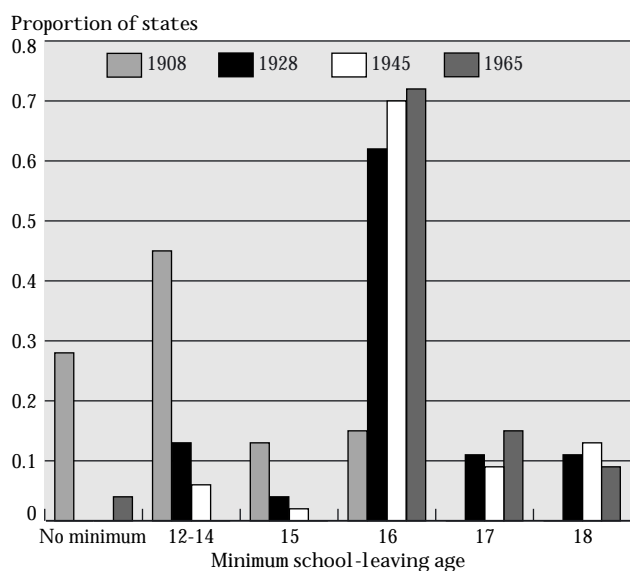
From an economist's perspective, the underlying goal of education is to prepare younger generations for success in adult life. A major determinant of adult success is earnings. Betts (1996a) surveys the literature that asks whether school spending affects students' earnings later in life, even if there appears to be little relationship between school resources and test scores. Quite a number of studies have found a relationship between adult males' earnings and school resources in their state of birth. But the literature is by no means unanimous. Work by Betts (1995a) and Grogger (1996), among others, shows that when school resources are measured at the school actually attended, the results are much more consistent with the test score literature: the impact of school inputs on earnings is not statistically significant.

More to the point, the estimated effect of raising school spending on students' subsequent earnings is extremely small. This is true whether one measures school resources at the school actually attended or the district attended, or whether one instead uses the person's state of birth to create a rough proxy for school resources.

The small impact of school spending on earnings stands in stark contrast to a voluminous literature documenting that a person's years of schooling are strongly related to subsequent earnings. Mincer's (1974) seminal contribution estimates that one additional year

Chart 2

Distribution of Minimum School-Leaving Ages across States, by Year



Source: Lang and Kropp (1986).

of schooling typically increases wages by about 7 to 10 percent. Psacharopoulos (1985), in a review of the literature, estimates that in developed countries a year of schooling typically is associated with a 9 percent rise in earnings. He also reports much higher returns to years of schooling in developing countries—a finding that invites us to ask what might be done in developed countries, such as the United States, to increase the payoff to a year of schooling.

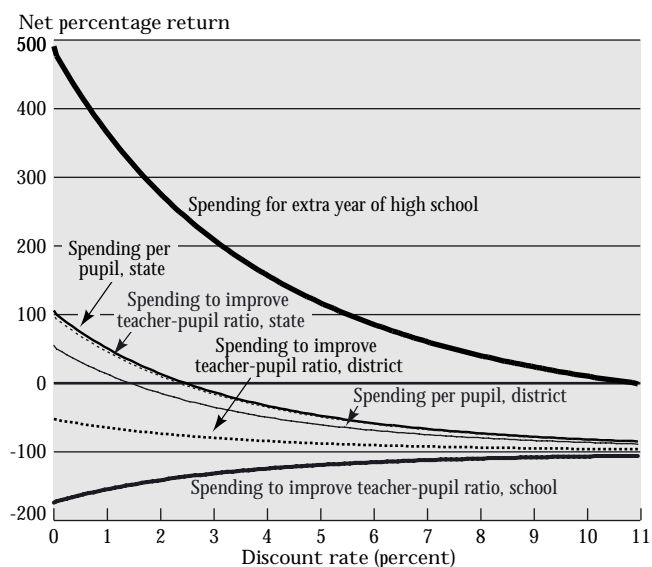
Simple calculations suggest that the economic returns to raising the school-leaving age are fairly substantial, but that the economic returns to increasing school spending are relatively meager. Consider first the rate of return to increasing spending per pupil or increasing the teacher-pupil ratio. It is important to discount future costs and benefits when calculating the net benefits from spending on a given school input, because all the costs are incurred in the early years of a person's life, while the benefits of higher wages accrue much later. The following calculations discount all costs and benefits to the year in which a student is in kindergarten. I assume that a 1 percent increase in spending per student is made in every school year between kindergarten and grade 12. To calculate the resulting increase in the discounted value of lifetime earnings, I multiply the predicted percentage wage gain, taken from the elasticities reported in Betts (1996a), by the discounted value of earnings for the average American male worker between the ages of nineteen and sixty-four. The calculation of the present discounted value of earnings uses the actual profile of earnings by age for male American workers, obtained using weighted earnings data from the March 1993 Current Population Survey tape. This tape contains information on annual earnings in 1992. The average annual earnings of male workers obtained from the Current Population Survey, taken as a simple mean across all ages from nineteen to sixty-four, was \$22,737.¹ Using this information, it is possible to calculate the net percentage return to an investment in school spending. This is calculated as the net return (wage gains minus the costs), divided by the costs, and expressed in percentage terms.

Similarly, one can calculate the net percentage return to an extra year spent in high school. Following Betts (1996a), assume conservatively that if we required a student who had dropped out of high school after grade 11 to remain in school for another year, the additional human capital he acquired would increase his earnings by 7.5 percent. This must be balanced against the cost of an extra year of schooling, which is wages forgone plus average spending per pupil in public schools, the latter of which is approximately \$5,000.² By calculating the estimated percentage wage gains and the initial wages lost from staying in school, again using the March 1993 survey, we can obtain different net returns on the initial investment for different "interest rates" or "discount rates."

The net percentage return to different types of educational expenditures is plotted in Chart 3. Note that as the discount rate rises, the net percentage return in general falls, because the given educational expenditure leads to higher predicted earnings for the student, but only later in life. As the discount rate rises, the present value of these wage gains declines.³

Chart 3

Net Percentage Return to Given Type of Spending versus Discount Rate



Source: Author's calculations, based on March 1993 Current Population Survey data.

The results are remarkable. The net percentage returns to requiring an extra year of high school are much higher than the returns to increasing school spending or to increasing the teacher-pupil ratio. As previously mentioned, when researchers have measured the actual school resources in a person's school district or school, the estimated effect of school spending is found to be smaller than when researchers instead use spending in the person's state of birth as a proxy. But even the most optimistic results—those that use the state-level data—suggest far smaller returns to increased spending per pupil or teacher-pupil ratios than are obtained when the student stays in school an extra year. Furthermore, in the district-level or school-level studies of the teacher-pupil ratio, even at a discount rate of 0 percent the initial costs of hiring more teachers are never recouped by subsequent increases in students' wages.

What is to explain the huge discrepancy between the returns to staying in school an extra year and the returns to increasing spending per pupil or increased teacher-pupil ratios? One possibility is that the returns to a year of education do not reflect true productivity gains. Instead, these returns might arise because of unobserved differences in ability between the highly educated and the less well educated. In other words, talented students might obtain more education merely to signal their ability to employers. If this "signaling" theory is true, then the observed returns to increasing the school-leaving age would be largely illusory.

Yet compiling evidence in favor of the idea that education merely signals a worker's productivity, rather than adding to productivity, has proved quite difficult. Lang and Kropp (1986) indirectly test for signaling. When a state raises its school-leaving age to, say, sixteen, then if students obtain education in order to signal their ability to employers, some seventeen- and eighteen-year-old students who would formerly have dropped out at age sixteen now begin to stay in school longer. They do this to "separate" themselves from their less able classmates. Lang and Kropp present some evidence that this has occurred historically. But numerous studies that have attempted to control for unobserved ability more directly have typically found that a year of schooling truly does increase productivity (see, for instance, Willis and Rosen [1979]). Numerous studies of

twins have attempted to sort out the true impact of a year of schooling on an individual's wage by comparing twins who obtained different levels of education. These studies have typically found that, if anything, the true productivity gain that results from an extra year of school is even higher than the 7.5 percent return that I used in the above analysis (see, for instance, Ashenfelter and Krueger [1994]). If these estimates are accurate, they suggest that increasing the school-leaving age would have substantial effects on the earnings of the affected students.

The best summary of the existing evidence is that, in the United States, the returns to increased spending per pupil or to increased teacher-pupil ratios have been extremely modest. The returns to an extra year of high school are much higher. However, as shown in Chart 3, even in this case the returns are not astronomically high: above a discount rate of 10.95 percent, the returns to additional education become negative. So, educational expenditures along the "extensive" margin (years of schooling) have been fairly productive, while expenditures along the "intensive" margin (spending per pupil) have had surprisingly small payoffs. This raises the question whether other aspects of public schools need to be changed in order to make financial inputs more effective than they have been in the past. The rest of this paper examines the proposition that the missing element in past reforms—the third leg of the stool—has been higher standards, higher expectations, and a higher degree of accountability in the nation's public schools.

A DEFINITION OF ACADEMIC STANDARDS AND ACADEMIC EXPECTATIONS, WITH EXAMPLES

It is easy to speak in general terms about higher standards and higher academic expectations for the nation's students. But what in practice does this mean? This section briefly outlines the necessary components for a variety of reforms that could represent a genuine shift toward higher standards.

CURRICULUM STANDARDS, TESTS OF ACHIEVEMENT, AND ACTIVE RESPONSES TO FAILURE

Perhaps the most often discussed way of increasing standards in schools is to strengthen the curriculum in

subjects such as mathematics, science, and English. As explained in the next section, many states have embarked on curriculum reform in one or more subjects. For curriculum reform to succeed, the subject content must be specified in detail, to ensure that all schools interpret the standard in the same way. A curriculum that states that “by grade 4, students will be able to express themselves well in written English” would not meet this criterion. It is so vague that it gives teachers little if any direction about what to teach. A far better approach would be to specify basic rules of grammar, spelling, and composition that should be mastered by that grade. The point of such a standard is not to tell teachers how to teach but to guide them as to what to teach, and when.

Creating a specific list of skills and knowledge that children should acquire in each grade is the first step. A curriculum standard by itself, however, is unlikely to improve schools substantially. It is also necessary to test students periodically to check whether they are meeting the standards set for each grade. Some states now have mandatory achievement tests in certain grades, but other states allow school districts to decide for themselves whether they want to test children. If so, the districts must choose among off-the-shelf tests or write a new test specific to their own curriculum. The need for a very specifically worded curriculum now becomes even more obvious: without a detailed curriculum, it will prove impossible to devise a test that gauges students’ academic progress.

The twin pillars of content standards, then, are a specifically worded curriculum and achievement tests that measure how well students are absorbing the prescribed curriculum.

It is certainly possible that testing based on content standards could improve the quality of schooling directly. Once parents become aware of any subject areas in which their children have fallen behind, they are likely to become more actively involved in their children’s schooling. Students themselves are likely to exert more effort during the academic year, knowing that at the end of the school year they will take a test that will inform their teachers and their parents about how much they have

learned. But there remains a distinct possibility that some students and some school administrations would disregard test results.

The investment in content standards and testing is likely to have a bigger payoff if there is something tangible at stake for both students and schools. The central question becomes, how, if at all, should a school react if a student does poorly on an achievement test? One obvious solution is for the school to devote additional resources—for example, tutoring or smaller classes—to such students. In other words, once a school system has developed a detailed curriculum and begins to test its students, it can direct additional spending to the students who need the most help. Systematic testing makes it possible to move away from a policy of improving schools through expensive systemwide increases in spending. This more focused approach to spending might make additional spending on schools much more effective than it has been in the past. To some extent, schools already direct more resources toward students whose achievement lags their grade level (see, for instance, Betts and Shkolnik [forthcoming]). However, the institution of a specific curriculum and regular testing based on this curriculum would allow more effective targeting of additional school expenditures toward children of relatively low achievement.

A second way in which schools can react when a student performs poorly on a test is to give the student a direct stake in his or her academic progress. A very long tradition in American schools has been to hold students back a year if they have not progressed sufficiently. Another approach is to require students who are lagging behind to attend summer school. Students will clearly want to avoid either of these outcomes, and thus will be motivated to work hard during the school year. The next section will discuss an innovative program recently implemented in Chicago that seeks to hold students accountable for their progress, yet gives failing students a second chance to exhibit their academic prowess.

The idea that testing would be more effective if something tangible were at stake is equally applicable to students and to their teachers and schools. Later on, I will discuss how school systems can use—and in some cases

already are using—students' test scores to identify schools that are failing, and some of the ways in which school administrations are intervening in such cases.

GRADUATION EXAMS AND EXIT EXAMS FOR SCHOOL LEAVERS

The above section argues in favor of regular monitoring of students as they progress through school. A closely related idea is to test students in grade 12 to ensure that they will graduate from high school with a skill level commensurate with the demands of the labor market. As noted in the next section, some states are moving in this direction and one—New York—has had a system similar to this one in place since the last century.

A less widely practiced policy is to provide a test of achievement that all students must pass before being allowed to graduate from high school *or to drop out of high school*. Such a test would ideally be offered to students in grade 9 or 10 and would focus on basic skills: reading, writing, and mathematics. In most states, such a policy would represent a sea change in how those likely to drop out of school are treated. In many states, students are allowed to drop out of school when they reach the age of sixteen, without having to demonstrate command of even the most basic skills. In today's economy—in which new technologies and changes in international trade patterns have acted to shift employers' needs toward more highly skilled labor—student dropouts have fared particularly badly over the last two decades. During the 1980s, the real wages of those with a college degree held fairly steady, while the real wages of those with a high school diploma or less fell substantially. For instance, Blackburn, Bloom, and Freeman (1990) report that between 1973 and 1987 the earnings of white male high school dropouts who worked full time and full year and were twenty-five to thirty-four years old fell from \$20,128 to just \$15,922. (Both of these figures are expressed in 1987 prices.) Such a precipitous fall in earnings over this short period suggests that schools should be particularly concerned about how well they are preparing their weakest students for the modern labor market. Since so many of these students ultimately drop out of high school, it stands to reason that the criterion for

dropping out should not simply be age, but should instead be a minimal level of achievement on a test of basic skills.

Another advantage of such a policy shift is that it gets the incentives right for students. A high school teacher will have little impact on a disaffected fifteen-year-old student who can feel free to ignore class assignments and so on, knowing that he or she can drop out at will after turning sixteen. By supplementing or replacing the policy of a minimum school-leaving *age* with a minimum school-leaving *level of achievement*, schools will give such students the right incentives to make the most of the time they do spend in school. Students will realize that if they cannot demonstrate mastery of a core set of skills—such as reading, writing, and basic mathematics—they will have to remain in school until they are able to do so. Of course, accommodation will be required for students with learning disabilities.

HIGHER GRADING STANDARDS

Another component of a school's overall standards is the way in which its teachers assign letter grades. If a school makes it overly easy to obtain a grade of A, the school is likely to reduce many students' effort. The best students, having obtained the top grade with little effort, will not find it worthwhile to work harder, because when they apply for a job or for admission to a university, their transcripts will not convey to potential employers this additional effort. For those who are not already A students, weaker grading standards can also be counterproductive in the sense that if any of these students think that a potential employer will care about letter grades, they can receive adequate letter grades without exerting much effort. Similarly, if B or C students hold any hope of attending a university, then letter grades should be of direct concern to them, so higher grading standards should induce additional effort.

HOMEWORK

The above suggestions deal with specific examples of how a school can set higher standards. Another step that a school could take is to set higher expectations of its students. A primary example of this is a school that encourages teachers

to assign more homework to students, especially in core subjects such as English and math. The next section provides a summary of a growing body of evidence supporting homework as a key to creating better schools.

THE EFFECTIVENESS OF HIGHER EDUCATIONAL STANDARDS IN PRACTICE

CURRICULUM STANDARDS AND TESTS OF ACHIEVEMENT: ARE THEY WIDELY USED? DO THEY WORK?

One of the main proponents of curriculum standards and achievement testing has been John Bishop. In a series of papers, Bishop (1996, 1997) has gathered indirect evidence that such standards can palpably improve the quality of education. For instance, he reports that students from the state of New York tend to outperform students from other states on standardized tests such as the Scholastic Aptitude Test (now known as the SAT). While there are literally dozens of potential explanations for such a finding, one is that New York has stood alone in setting a statewide exam for high school seniors: students in New York have long had the option of taking the Regents examinations. Those who pass the exams receive a high school diploma different from that received by students who opt not to take the Regents exams. Passing the exams has in general been a prerequisite for college entry in New York.

New York is now in the middle of a quite bold experiment, in which the alternative high school diploma, known as the local diploma, is being phased out. Current plans are for the Regents exams to become a requirement for high school graduation for all students in the state by 2003 (*New York Times* 1996).

Bishop (1996, 1997) has also analyzed the performance of Canadian students on the 1991 International Assessment of Educational Progress (IAEP). Canada's educational system is in many ways similar to that of the United States, but significantly, many of Canada's ten provinces now require students to pass a provincewide exam before graduating from high school. Bishop finds that students from the provinces that have implemented graduation exams tend to perform significantly better on

the IAEP. A clear concern in all empirical research that uses differences in policies across political borders is that variations in the given policy might be endogenous. That problem appears especially likely in this context: for political reasons, a ministry of education is less likely to institute provincewide testing if it knows that its students are likely to fare poorly. Since research consistently finds that the socioeconomic background of parents is highly predictive of students' achievement, it could be that only richer provinces would institute testing in the first place. However, the provinces that have instituted testing include not only the most economically developed, but also the least economically developed: Newfoundland. This suggests that the reported correlation between the existence of graduation exams and student test scores reflects true causation. Bishop (1996) also reports corroborating facts. For instance, he finds that in provinces that have instituted graduation exams, students report watching less television.

In the United States, how widespread is the idea of graduation exams across states? New York has offered the Regents examinations to high school students for over a century. But until 2003, these exams will be partly voluntary; students can instead opt for the local diplomas that do not require the test. In many other states, work is under way to develop curriculum standards, and in a subset of these states plans are also under way to require high school seniors to write graduation exams that are linked to the curriculum.

The American Federation of Teachers (AFT) has recently committed to performing an annual evaluation of states' progress in developing precisely worded curriculum standards and tests. Their 1996 report suggests a crazy quilt of reforms across the states. The AFT finds that all states apart from Rhode Island and Wyoming are now developing grade-by-grade content standards. Unfortunately, there is an extraordinary disparity in the level of detail provided in each of the state curricula. The AFT reports that only Virginia has developed curriculum standards in English, math, science, and social studies that are sufficiently clear and explicit to provide guidance to teachers and parents about what should be taught in each grade. Of the forty-eight states that are writing or have

written curriculum standards in at least one of these subjects, forty-two have developed or are developing tests that will be based on the content of the curriculum. However, only fifteen of these forty-two states have planned or are planning test programs in all four subject areas mentioned above, with tests being based on precisely worded content standards. In summary, most states are now developing content standards in at least a few of the key subject areas and are developing statewide tests of these curricula. But only a minority has as yet developed a comprehensive set of content standards backed by testing.

ACHIEVEMENT TESTS, REMEDIATION, AND GRADE RETENTION POLICIES

Earlier, I argued that a specifically worded curriculum backed by periodic testing of students could by itself significantly improve school quality, because it provides an objective report card on individual student achievement. When a student falls behind grade level, the student, his or her parents, and the school's teachers can react. But it seems reasonable that content standards and testing will be more effective if school systems have a formal plan in place to deal with students falling behind grade level. Another critical question is whether the student has anything at stake. Will a student who is far behind grade level have to attend after-school tutorials, summer school, or—in extreme cases—even repeat the same grade next year? Alternatively, do students not take the tests seriously, because there are no consequences attached to poor performance on them? And if the school system does mandate remediation, has it put in place a funding mechanism?

The AFT (1996) report provides partial answers to these questions. Its survey reveals that eighteen of the fifty-one states (the District of Columbia is treated as a state) have state-mandated remediation for students who do not meet the state's educational standards. However, only ten of these eighteen states also provide funding to schools for the additional teaching.

What are the state-mandated consequences for students who fail to do well on the achievement tests? The AFT study finds that even though forty-two states have planned or are planning statewide tests, only a handful

have made promotion between grades or graduation from high school conditional upon test scores. The more common action, adopted or about to be adopted by thirteen states, has been to institute graduation exams that test whether a grade 12 student achieves at a level equivalent to grade 10 standards or higher. This is a noteworthy trend.

However, it is surprising how few states have made student promotion decisions based in any way on students' results on the state tests. The only states or districts to have done so by 1996 were the District of Columbia, North Carolina, and South Carolina.

Why have only three of the forty-two states with tests and curriculum standards linked test performance to grade promotion? One explanation is that the empirical literature on the consequences of grade retention has in general found that holding a student back one year has either zero effect or a negative effect on the student's subsequent rate of learning. For instance, in a review of the literature, Holmes (1989) concludes that most studies have found that grade retention is associated with poorer performance after the student is held back a year. Only nine of sixty-three studies found that retention improved student performance. Holmes indicates that in most of these positive studies the "treatment" of students was not simply retention but retention accompanied by quite intensive remediation. It is therefore not clear whether students who were held back a year did better than the comparison group because of the additional year or the extra help they received.

The above summary of state policies on curriculum standards and testing suggests that most states are now working toward these goals, but that in most cases much remains to be done. It also highlights the substantial diversity across states in their policies concerning educational standards.

The state-level summary ignores the fact that in most states individual school districts enjoy considerable autonomy to create their own programs to supplement or strengthen statewide initiatives. For this reason, the above summary is likely to understate the extent to which students in public schools are held to curriculum-based standards. Similarly, the summary is also likely to understate the disparities across schools in the stringency of

academic standards. Interschool and interdistrict variations in curriculum content and in standards are of particular concern given that each year so many children switch schools when their parents move. For this reason, it would seem worthwhile for a state to set high standards in order to level out any existing variations in standards across districts.

A detailed analysis of how individual school districts within even one state set and enforce standards would be a major undertaking. But certain school districts around the country have received national attention for their innovations in setting academic standards. The next section discusses a particularly bold set of reforms that the Chicago public schools have recently implemented.

THE CHICAGO PUBLIC SCHOOLS EXPERIMENT

The Chicago public schools (CPS) system is one of the largest school districts in the country. Its students represent an ethnically and racially diverse group that must contend with all the challenges of life in a modern urban area. It is perhaps not surprising, then, that in Chicago students' academic performance has long lagged behind national averages. For example, throughout the 1990s grade 9 students' average performance on the Tests of Achievement and Proficiency (TAP) has hovered between the twenty-third and thirty-sixth percentiles of national norms. Results for other grades have fallen into a similar range.⁴

The CPS has given tests to children in various grades throughout the 1980s and 1990s. But during the 1996-97 school year, school administrators made the tests "matter." In grades 3, 6, 8, and 9, students whose performance lagged behind national norms on either the reading or mathematics portion of the tests were required to attend summer school. The cutoff points below which students were required to attend summer school were 2.8 for grade 3, 5.2 for grade 6, 6.8 for grade 8, and 7.9 for grade 9. (The tests were given in the spring, so that a student progressing at the normal rate would be at grade level at the time of the test, while a score of 2.8 in grade 3 would suggest that the student was approximately two months behind national norms.) The summer school, known as the Summer Bridge Program, lasted for six to seven weeks. At the end of the Summer Bridge Program, students took the tests a second

time—the TAP in grades 8 and 9 or the Iowa Tests of Basic Skills in lower grades. A student who met the cutoff grade equivalents listed above for both reading and math by the end of summer was allowed to advance to the next grade.⁵ Students who did not reach the cutoff level in either test were required to repeat the grade.⁶ During the 1997-98 school year, students who were held back were in many cases to receive additional help, typically in the form of tutorial classes.

Results from the 1996-97 school year are quite remarkable, both for the sheer number of students who failed the initial tests in spring, and for the sizable gains in achievement recorded for those who entered the Summer Bridge Program. Table 1 reports the number of students who failed to reach the cutoff in either math or reading in

Table 1
RESULTS FROM THE CHICAGO PUBLIC SCHOOLS' TESTING AND SUMMER BRIDGE PROGRAMS

Variable	Grade 3	Grade 6	Grade 8	Grade 9
SPRING TEST RESULTS				
Students taking math test	23,989	25,275	22,708	22,986
Students taking reading test	24,124	25,311	22,776	22,967
Students taking at least one test	24,124	25,311	22,776	22,986
Students failing at least one spring test	11,632	8,870	6,180	14,287
As a percentage of students taking at least one spring test	48.2	35.0	27.1	62.2
SUMMER PROGRAM AND SUMMER TEST RESULTS				
Students taking summer test	10,336	8,275	5,831	9,610
As a percentage of students who should have taken summer test	88.9	93.3	94.4	67.3
Students passing summer test	4,236	3,668	2,891	3,695
As a percentage of students taking summer test	41.0	44.3	49.6	38.4
As a percentage of students who should have taken summer test	36.4	41.4	46.8	25.9
Percentage of all students taking spring tests who passed by end of summer	69.3	79.4	85.6	53.9
GAINS IN PERFORMANCE AMONG SUMMER STUDENTS				
MEAN GRADE EQUIVALENT (GE) OF STUDENTS				
Spring GE of students below promotion level on spring reading test	2.03	4.41	5.88	6.32
August GE of students below promotion level on spring reading test	2.47	5.11	6.87	7.79
Gain in reading GE over summer	0.44	0.70	0.99	1.47
Spring GE of students below promotion level on spring math test	2.29	4.58	6.25	6.69
August GE of students below promotion level on spring math test	2.98	5.20	7.04	8.03
Gain in math GE over summer	0.69	0.62	0.79	1.34

the spring tests, both as a raw number and as a proportion of all CPS students who took the spring tests at that grade level.⁷ It shows that 27.1 to 62.2 percent of students failed at least one of the two tests, depending on the grade level.

The table also reports outcomes of the Summer Bridge Program. There are two complications that must be dealt with. First, I exclude from the analysis students who had met the cutoff scores in the spring test but who enrolled in the Summer Bridge Program on the advice of the school because they were near the margin. Such students automatically “pass” the test in August. The second complication derives from the fact that only a fraction of the students who should have enrolled in the Summer Bridge Program took the tests at the end of summer. For this reason, I calculate pass rates at the end of summer using in the denominator both the total number of students who took the summer test and the total number who *should* have taken the summer test. Using the latter number, I find that 25.9 to 46.8 percent of students who should have taken the summer test passed. Note that a substantial fraction of students in grade 9 did not take the summer test. The most likely explanation for this is that by CPS policy, students who were fifteen by December 1, 1997, were not to be held back in grade 9. When the pass rates are calculated as a percentage of those who actually took the summer tests, the success rate is much higher, ranging from 38.4 to 49.6 percent, with the highest success rate observed among grade 8 students. But overall, well over half of the regular students in these grades are promoted at the end of the school year. The highest success rate is 85.6 percent, among grade 8 students.

Clearly, the Summer Bridge Program is not a panacea for students who initially obtain low scores: over half of those enrolling in the program do not pass in their second attempt. But on average, students progressed remarkably during the program. Table 1 also shows the initial mean grade equivalent of those who failed to meet the promotion criterion in the given subject, and the mean grade equivalent that these same students obtained after the Summer Bridge Program. The mean gain in grade equivalents is typically one half year to a full year or even

more. Given the low base from which these students began in the spring, they remain on average one-half grade to one grade equivalent behind by August, or in the cases of reading scores in grades 8 and 9, slightly more than one grade equivalent behind. The observed improvements, however, are extremely impressive.

Of course, one concern raised by these data is that we are observing “regression to the mean.” If a student has a bad day when taking the spring test and scores considerably below his or her potential, that student might have to enroll in the Summer Bridge Program. In such a situation, the student might score much higher on the second test in August, not because of the usefulness of the program but because the low test score in the spring was a statistical aberration. It would be useful to study patterns in the test scores of students *prior* to the year they are required to attend summer school, to determine whether some of the summer school students were forced to attend simply because of a bad day in the spring test. Similarly, it would be useful to follow these students for at least one more year to check whether the remarkable gains in achievement over the summer endure.

Nevertheless, the fact remains that nearly half of Chicago’s summer school students had by the end of the summer improved sufficiently to meet the promotion criterion. Improvement on this scale suggests that the Chicago public schools system has found an ideal incentive system for students. The impressive gains in mean grade equivalent, of anywhere from 0.6 to 1.3 years in the space of the six- to seven-week Summer Bridge Program, point in the same direction. Low academic achievement now has tangible consequences for students: the students must spend extra time in remedial classes, both in the summer and during the school year. The CPS has enforced reasonably high standards and gotten the incentives right at the same time that it has directed additional funding toward students whose achievement has lagged the most. Over time, this fledgling program could do much to improve the academic achievement of students who are most in need.

The CPS program also improves on the traditional “fix” for education, in which wholesale increases are made in school spending. By identifying students most in need

through tests each spring, scarce financial resources are being targeted toward students who truly need additional time with teachers.

It is possible to compare at least roughly the costs and benefits of regular schooling with those of the summer school program. A typical American school spends about \$5,000 per pupil per year, and the typical student in that school will gain one grade equivalent over the school year. In 1997, the CPS spent about \$34 million on its Summer Bridge Program, or about \$720 per participant. This sum translates into about \$1,000 for each student who took the test in August, since some marginal students who participated in the Summer Bridge Program voluntarily did not take the test at the end of summer. This latter figure represents about one-fifth the cost of a year of regular schooling for each student who participated. Yet the gain in achievement far surpasses one-fifth of a grade equivalent. As shown in Table 1, the mean gain over the six- to seven-week summer program ranged from a low of 0.44 of a grade equivalent in grade 3 reading tests to 1.47 grade equivalents in grade 9 reading tests. Even assuming that these students gained a full grade equivalent during the regular academic year, on a dollar-for-dollar basis the Summer Bridge Program is anywhere from 2.2 times to 7.4 times as effective as schooling during the regular year. By any standard, these additional expenditures appear to have been much more effective than traditional expenditures made during the academic year. Over time, it will become possible to test whether these gains are permanent or transitory.

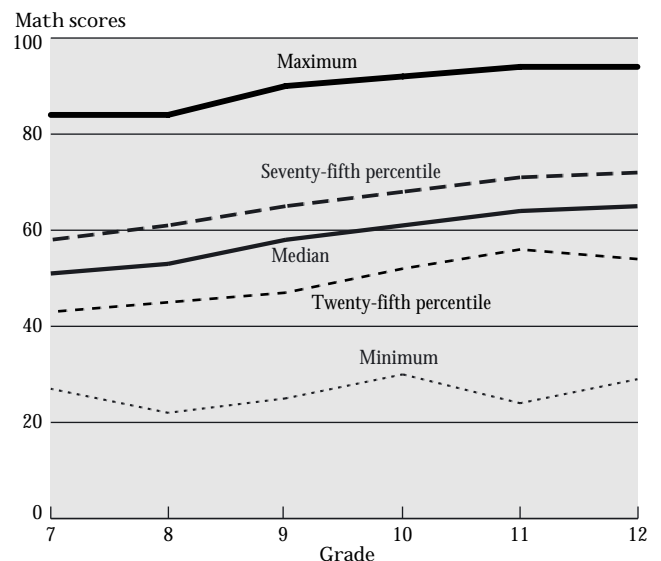
Some readers will rightly wonder whether the initial failure of a number of students to meet the promotion criteria simply reflects adverse conditions in the Chicago public schools relative to those in public schools elsewhere. Consider the following simulation. Suppose that national standards were put in place dictating that any secondary school student more than a year behind in mathematics achievement was required to attend summer school. Suppose that the criterion for being “more than a year behind” was that a student in one grade had a test score below the median test score of students in the previous grade. What would happen?

Chart 4 depicts the median math test score in a representative sample of American high school students by grade level, which I calculated from the Longitudinal Study of American Youth (LSAY). (This study followed approximately 6,000 students over a five-year period.) The chart also shows the test scores of students at the twenty-fifth and seventy-fifth percentiles, and the minimum and maximum scores observed in each grade. The most striking aspect of the chart is the huge disparity in achievement among students within any grade level. In the sample, the increase in the median math test score between grades is approximately 2.8 points out of 100. This median gain is tiny compared with the dispersion within any grade: the gap between the test scores of students in the twenty-fifth percentile and the scores of students in the seventy-fifth percentile within a grade is typically about 16 points.

The slow growth in student performance *across* grades relative to the large degree of dispersion *within* grades in this nationally representative data set suggests that Chicago’s experience is not atypical. The LSAY data clearly indicate that large proportions of students lag

Chart 4

Math Test Scores by Grade in the Longitudinal Study of American Youth



Source: Author's calculations, based on data from the Longitudinal Study of American Youth.

Table 2
**PERCENTAGE OF STUDENTS WHOSE MATH SCORES
 WOULD MISS CUTOFFS BASED ON MEDIAN SCORES
 IN PRIOR GRADES**

Grade	Percentage below Median Score in Previous Grade	Percentage below Median Score Two Grades Earlier
8	40.8	N/A
9	36.7	31.6
10	39.0	26.4
11	39.1	29.2
12	45.5	39.7

Source: Author's calculations, based on math test score data for students participating in the Longitudinal Study of American Youth.

behind national norms in schools across the country. Table 2 shows the percentage of students in the LSAY who would have to repeat a grade if the criterion for promotion was that a student in grade “N” needed a test score equal to or above the median score among students in grade “N-1.” The percentages of students who would be required to repeat each grade are extremely high, and quite similar to what has been reported in Chicago, where the promotion criterion is that students must score, roughly speaking, within a year of the national norm for their grade level. The table also shows the percentage of students who would have to attend summer school under lower standards. Even when the promotion criterion is reduced so that a student needs only to obtain the median test score of students *two* grades back, anywhere from 26 percent of students in grade 10 to 40 percent of students in grade 12 would miss the cutoff point.

This simulation illustrates the extent of the problem in American public schools: disparity in achievement within grades is simply huge. It also shows that when school districts set reasonably high standards, they should be prepared for a large proportion of their students to fall below the standard, at least initially.

GRADING STANDARDS

A little-examined characteristic of schools is the stringency with which they grade students. If a school makes it easier to obtain a letter grade of A or B, students might respond by exerting less effort. Such a response will occur if students care directly about letter grades, as opposed to their true level of academic achievement. There are several

plausible reasons for this. First, students may care about letter grades simply because their parents care. Second, students may realize that employers may use letter grades as a signal of a student’s achievement, willingness to work hard, and so on. Some employers may examine high school transcripts directly, but it seems likely that an indirect mechanism is at least as important: firms rely on the postsecondary sector to identify high school students who have done well. College admission offices in effect act as an information intermediary between students and employers by closely examining high school transcripts in an effort to identify the students most likely to gain from further education.

In Betts (1997), I examine grading policies in math and science courses in a representative sample of American schools. By comparing students’ test scores in these subjects with their letter grades, I construct measures of the grading standards at each school. I find that American high schools differ radically in the way in which they assign letter grades to students of given achievement. I also find that the stringency with which a school grades is strongly related to the rate at which students learn. Even after controlling for the initial level of achievement of students at the school, traits of the individual student, traits of his or her family and peers, and detailed traits of the classroom, I observe that the school’s grading standard remains a highly significant and positive predictor of gains in test scores. Unfortunately, a policy of higher standards does not improve the performance of all students identically. Although C students benefit from attending a more rigorous school, A students benefit even more. The lesson is clear: in this instance, a policy of higher standards will also induce higher disparities in achievement. If a school administration is concerned with the distribution of student achievement, and not just the mean level of achievement, remedial policies to assist students with lower grade-point averages are in order.

HOMEWORK

Each of the above suggestions for improving schools involves setting higher standards. A closely related suggestion is to increase the amount of homework that

teachers assign. This is perhaps better thought of as a way of setting higher academic expectations than as a way of setting higher academic standards, although of course teachers can hold students accountable if they routinely fail to complete their homework. A policy of assigning more homework is quite similar in spirit to the earlier suggestions, because it too recognizes that one of the most important inputs in the “education production function” is the student’s own effort. This simple fact has been ignored in most of the traditional studies of whether “spending matters.”

A number of studies of the impact of homework on achievement have been carried out. Cooper (1989) gives a good review of the existing evidence. Two experimental studies have been performed, with somewhat mixed results. In both cases, the number of students involved in the experiment was very small, ranging from roughly 90 to 350 students. Cooper also reports that a number of correlational studies find a positive and significant link between the time students spend on homework and their achievement. The effectiveness of homework appears to be higher in secondary schools than in the elementary grades. Unfortunately, virtually all of the correlational studies test for a relationship between the amount of homework that students report doing and their achievement. This approach leads to a clear possibility of reverse causation: if better students routinely *choose* to do more homework than their peers, then the observed relationship between test scores and homework might be spurious.

A partial solution to this problem is to model student achievement not as a function of the homework that students do but as a function of the homework that teachers assign. In Betts (1996b), I use the aforementioned LSAY data for this purpose. I find a strong positive link between the amount of homework that teachers assign and the rate at which the student’s test score rises. I also estimate a “fixed-effect” model in which I use *variations* in the amount of homework assigned to individual students across grades to identify the effectiveness of homework, with similar results. The estimated effects of additional homework are quite striking. Using published estimates of the relationship between math test scores and

earnings, I calculate that an hour spent doing homework is equivalent to earning about \$6 to \$12. I make these estimates by discounting the future wage gains at a rate of 3 percent and by setting the opportunity cost of doing homework at the average earnings per hour of teenagers. Additional homework appears to be a particularly cost-effective method of improving school quality: it is the total amount of homework assigned—rather than the amount assigned, graded, and returned—that is more closely related to students’ rate of learning. The effectiveness of homework, by any measure, is quite large. The results suggest that among the students in grades 7 through 12 in the sample, a one-hour homework assignment is as effective as an hour spent in class.

HIGHER STANDARDS ARE NOT ONLY FOR STUDENTS: PROVIDING INCENTIVES TO TEACHERS AND SCHOOL ADMINISTRATORS

To this point, I have focused on the methods of improving incentives for students that target additional spending toward the students most in need. However, the same principle of tying additional spending to the setting of higher standards can be applied to teachers and entire school systems just as easily as it can be applied to students. Space constraints prevent me from developing this theme in detail, but the following discussion highlights the main arguments.

The essential point is that there are good teachers and bad teachers, effective principals and less effective principals. What, then, should a school board do when, after setting higher standards for its students, it realizes that at some schools the only remedy for low achievement is to improve the quality of teaching?

Improving teacher quality requires a two-pronged approach—setting up the economic incentives required to attract well-trained college graduates to the teaching profession while providing opportunities for more experienced teachers to gain new skills. Typically, school districts set teachers’ salaries as a fairly rigid function of teachers’ years of experience and the degrees they hold. To obtain a permanent certification, teachers in many states must obtain a certain number of graduate-level credit hours, which

typically lead to a Master's degree. Yet the evidence that teachers who hold a Master's degree are better teachers is decidedly mixed (see, for instance, Betts [1995a], Grogger [1996], and Betts [1996a]). Paying teachers by the level of degree held may make less sense than paying teachers extra for any college courses that pertain to their field of teaching in the school. A number of authors—for instance, Goldhaber and Brewer (1997) and Betts (1995b)—have found that college training in the field taught is related to teaching ability. Similarly, mechanically linking teachers' pay to years taught may not be the optimal policy: evidence suggests that teachers, especially after the first few years of teaching, do not necessarily continue to improve their quality of teaching much over time. Similarly, it is important for school administrators to respond to market forces. Murnane et al. (1991) show that over the last twenty years the starting salary for teachers has consistently lagged behind that of college graduates who work in industry. They argue strongly in favor of merit pay for teachers as a method of ensuring that the best teachers remain in the profession.⁸

The question immediately arises, how can principals identify the best teachers in order to allocate merit pay? The need for ongoing and objective assessment of teachers provides yet another reason why it is so important for school districts to set out a clear curriculum and then to assess students. But it would be wrong to reward teachers solely on the basis of the mean test scores in their classes. Given evidence that family background and peer effects strongly influence student achievement, such a policy would in effect punish teachers who taught in disadvantaged neighborhoods. Within schools, it would aggravate any tendency that might already exist for teachers to prefer to teach the most advanced students within each grade. A more reasonable approach might be to establish merit pay as a function of how much student achievement *improves* over one or two years.

Similar policies of merit pay for principals might also work. A number of school districts around the country have gone further, setting the pay of superintendents of entire districts to reflect the rate of improvement of the district's students.

Merit pay, additional course work for experienced teachers, and a less rigid structure for teacher salaries that would allow schools to attract talented college graduates in fields such as math and science are all good ideas. In some cases, however, more radical solutions might be needed. Perhaps in recognition of this, a number of state legislatures over the last few years have attempted to remove teacher tenure, so that school districts would find it easier to fire teachers who were not performing adequately. To the best of my knowledge, none of these reforms has met with success, because of opposition from a number of sources, including—not surprisingly—teachers' unions. Individual school districts have also started to put on "probation" those schools whose students fare poorly, and in some cases have "reconstituted" entire schools by firing or reassigning virtually all employees, from the principal down to the custodial staff, in a bid to change the prevailing culture at the school. It is too early to know whether such radical restructuring has had the intended results.

A different method for improving schools is to increase the degree of competition between schools. Basic economic models argue that when a firm has a monopoly, it will restrict supply and charge higher prices than it would if there were a high degree of competition in the market. By analogy, when a school district is the sole provider of schooling, the lack of competition allows it to do less with each education dollar than it would under competition. Ballou (1996) provides a recent and interesting example of how a lack of competition may render public schools inefficient. He finds that when hiring teachers, public schools do not seem to give any preference to applicants who have superior academic records. This finding is puzzling, given evidence by Ehrenberg, Goldhaber, and Brewer (1995) that cognitive abilities of teachers are positively and significantly related to the rate at which their students learn. Ballou's conclusion is that public schools face little competition for students, and so do not invest sufficient effort in finding the best applicants for teaching jobs. This conclusion, if true, would be an example of an inefficiency that would surely disappear if schools competed with each other more strongly for students.

A companion paper in this volume by Hoxby (1998) describes in detail the evidence that competition—whether between school districts or between public and private schools—might improve the efficiency of public schools. In addition, the paper by Rouse (1998) in this volume addresses the effectiveness of vouchers.

While the argument in favor of increased school choice appears to have some empirical backing, it is crucial to understand that such choice can work only if parents make informed decisions. Informed decisions require good information about schools. This provides yet another rationale for regular testing of students. Without a districtwide—or, preferably, city- or statewide—report card on student achievement and gains in student achievement at each school, it is unlikely that parents will be able to make informed decisions about the schools that are best for their children.⁹

OBSTACLES TO HIGHER STANDARDS AND SOME SUGGESTED SOLUTIONS

If academic standards—in the form of a clearly worded curriculum and tests designed around the curriculum—are such good ideas, then why do we not see more widespread use of these tools? Critics have raised many specific objections, but it is crucial to realize at the outset that testing threatens many interests. What politician wants to have it publicized that schools within his or her district are not adequately serving students? Testing can also cause discomfort for teachers, students, and in some cases parents. Because assessment, done properly, provides an objective “report card” on schools, it often provides impetus for radical change. Many professionals in education will naturally resist change.

Existing theoretical work on educational standards points to a second reason for opposition to higher standards. Two models developed by Costrell (1994) and Betts (1998) differ in some regards, but both establish that whenever a school raises its standards (such as a pass-fail standard), some students will lose out. In essence, a student whose ability or diligence was such that he or she was initially indifferent to meeting the standard or falling below it will choose not to exert the extra effort required after the standard is raised. This can lead to a significant

drop in well-being for such students. This theoretical result suggests that in the real world, school administrators who raise standards will have to devote additional attention to “borderline” students to ensure that they are able to continue meeting the requirements for grade promotion or for high school graduation.

A commonly heard complaint about testing of curriculum standards is that it will lead to teachers wasting time “teaching to the test.” Why should teachers be forced to squander valuable class time helping children to memorize facts and tricks for a multiple-choice test when they should be steering children toward more profound forms of learning, or so the argument goes. There is no doubt an element of truth in this statement. But, ideally, a test should not only require regurgitation of memorized facts, but should also assess a student’s ability to synthesize, to apply concepts learned in one context in a new environment, and so on. Ideally, then, tests will include not only multiple-choice questions but also questions requiring a written response. In short, if administrators write a test properly, teaching to the test is exactly what teachers should be doing.

Recent experience indicates that the main barriers to higher educational standards backed by testing are political. President Clinton’s call for voluntary national tests in reading in grade 4 and in math in grade 8 has recently met strong opposition on Capitol Hill.¹⁰ Some legislators have objected that federally backed tests represent an intrusion by Washington into education, which traditionally has been controlled at the state level. President Clinton has responded, with some justification, that the basic elements of mathematics are the same regardless of the state in which the student in question lives. Why should children—and schools—in some states be held to standards lower than those in other states?

One can also make a case for national standards on cost grounds. Surely, a reasonable set of national standards could be drafted at far lower expense than could fifty sets of standards, each specific to a given state. The problem is compounded by the actions of the many school districts that have drafted their own content standards in recent years. Not only does this lead to needless duplication of

effort, but it can create problems for students who move between school districts within a state at some time.

Clearly, however, when representatives from fifty states meet to attempt to establish national standards, diverse opinions are likely to lead to diluted national standards. Fear of such an outcome may explain why so many states have taken it upon themselves to develop curriculum standards, and why many school districts have developed their own content standards to supplement those provided by the state. It would appear that proponents of national standards will have the greatest chance of success if they focus on subjects in which there is general agreement about required elements of a core curriculum. For instance, the relative success of the National Council of Teachers of Mathematics in its attempt to create national standards in math stands in strong contrast to the recent failure of an attempt to create national standards in the much more contentious subject area of history. The divergence in outcomes may reflect underlying disagreements about what is important in an area. Therefore, it might be advisable, at least at first, for educators seeking national or even state-level standards to focus on areas such as mathematics, reading, and writing.

Space constraints prevent a further discussion of barriers to higher standards. However, readers interested in these issues should see Ravitch (1995) for a compelling insider's account of the history of educational standards in the United States. In particular, Chapters 5 and 6 of her book develop a clear prescription for change that is appealing in an economic sense, yet takes heed of the political realities of school reform today.

CONCLUSION: A CHECKLIST FOR REFORM

In the 1990s, virtually all states have started to develop curriculum standards to increase student achievement. Although the states vary remarkably in the number of subject areas for which they are developing curriculum standards—and in the specific content of the standards—as a group they have clearly made significant progress.

A necessary companion policy to higher standards is increased spending on assessment of student achieve-

ment. Testing is crucial. To ensure that reforms are effective, administrators must do more than prescribe a minimal curriculum in key subjects; they must evaluate the extent to which individual students meet the standards. Such tests serve a dual purpose: not only do they create an incentive for students to exert effort, they also provide a means to make teachers, schools, and entire school systems accountable to the public. The regular publication of test results by school is also an indispensable tool if public schools introduce a system in which parents can choose the school their children attend.

In this paper, I have reviewed recent attempts by states and school boards to raise standards. Limited empirical evidence suggests that higher expectations—whether established through higher graduation standards, more stringent homework requirements, higher grading standards, or increased requirements for promotion between grades—can spur student achievement. The paper also reviews the serious roadblocks that have hindered attempts to tighten standards, and suggests some ways in which school administrators can address the concerns of critics of testing and standards.

The two most historically important reforms to public schools in this century—raising the school-leaving age and increasing spending per pupil—in a sense form a two-legged stool. Allowing students to leave school at a certain age without having them demonstrate a minimal level of achievement is a shortsighted policy. Achievement, not age alone, should determine when a person is ready to leave school. Similarly, large increases in school spending that are not accompanied by increases in the standards to be met by students and their schools are likely to achieve little. Only by coupling these two policies to higher standards—and by testing the ability of students to meet the standards—are we likely to see large improvements in school quality. For this reason, the reforms in educational standards currently under way in many states have the potential to be surprisingly more effective than previous reforms.

ENDNOTES

The author thanks Sandra Storey and Joseph Hahn for helpful discussions concerning the Chicago data.

1. Male workers are chosen, since the studies reviewed in Betts (1996a) examine earnings of men only.
2. In 1990-91, total spending per pupil in American public schools was \$5,320, while the current expenditure was \$4,847 per pupil (National Center for Education Statistics 1991, p. 155).
3. The one exception is for estimates of the return to increasing the teacher-pupil ratio derived from school-level studies. Betts (1996a) finds that the average effect of this intervention is actually slightly negative, so that increasing the discount rate lowers the predicted losses.
4. All information in this section that relates to the Chicago public schools was obtained directly from the CPS system, except where noted.
5. A student who passed one test in the spring and the other test at the end of summer school was promoted. In addition, a small number of students who met the grade equivalent criterion for both reading and mathematics in the spring, but whose scores were borderline, voluntarily enrolled in the Summer Bridge Program.
6. The consequences for grade 9 students whose test scores were too low were slightly different. Any such student who was at least fifteen years old by December 1, 1997, was not retained, but was sent to a special remedial school, known as a High School Transition Center.
7. Students enrolled in the bilingual education program are exempted from participation in the Summer Bridge Program for up to three years. Similarly, special education students are not required to meet the standards. Accordingly, all calculations in Table 1 exclude these two types of students.
8. See, however, Chapter 5 of Ballou and Podgursky (1997), which argues that attempts to strengthen the incentive structure faced by teachers are unlikely to have much success, given the likelihood of opposition from entrenched interests.
9. See Hanushek et al. (1994, Chap. 6) for a detailed summary of ideas on how administrators could use incentives to improve schools.
10. See, for instance, Applebome (1997).

REFERENCES

- American Federation of Teachers*. 1996. MAKING STANDARDS MATTER, 1996: AN ANNUAL FIFTY-STATE REPORT ON EFFORTS TO RAISE ACADEMIC STANDARDS. Washington, D.C.
- Applebome, Peter*. 1997. "President's Proposal for National Reading and Math Tests Is Facing Wide Opposition." NEW YORK TIMES, August 3: 14.
- Ashenfelter, Orley, and Alan B. Krueger*. 1994. "Estimates of the Economic Return to Schooling from a New Sample of Twins." AMERICAN ECONOMIC REVIEW 84: 1157-73.
- Ballou, Dale*. 1996. "Do Public Schools Hire the Best Applicants?" QUARTERLY JOURNAL OF ECONOMICS 111: 97-133.
- Ballou, Dale, and Michael Podgursky*. 1997. TEACHER PAY AND TEACHER QUALITY. Kalamazoo, Mich.: W.E. Upjohn Institute for Employment Research.
- Betts, Julian R.* 1995a. "Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth." REVIEW OF ECONOMICS AND STATISTICS 77: 231-50.
- _____. 1995b. "Which Types of Public School Spending Are Most Effective? New Evidence on the School Quality Debate." University of California, San Diego, Department of Economics Discussion Paper no. 95-03.
- _____. 1996a. "Is There a Link between School Inputs and Earnings? Fresh Scrutiny of an Old Literature." In Gary Burtless, ed., DOES MONEY MATTER? THE EFFECT OF SCHOOL RESOURCES ON STUDENT ACHIEVEMENT AND ADULT SUCCESS, 141-91. Washington, D.C.: Brookings Institution.
- _____. 1996b. "The Role of Homework in Improving School Quality." University of California, San Diego, Department of Economics Discussion Paper no. 96-16.
- _____. 1997. "Do Grading Standards Affect the Incentive to Learn?" University of California, San Diego, Department of Economics Discussion Paper no. 97-22.
- _____. 1998. "The Impact of Educational Standards on the Level and Distribution of Earnings." AMERICAN ECONOMIC REVIEW 88, March.
- Betts, Julian R., and Jamie L. Shkolnik*. Forthcoming. "The Effects of Ability Grouping on Student Math Achievement and Resource Allocation in Secondary Schools." ECONOMICS OF EDUCATION REVIEW.
- Bishop, John H.* 1996. "The Impact of Curriculum-Based External Examinations on School Priorities and Student Learning." INTERNATIONAL JOURNAL OF EDUCATION RESEARCH 23: 653-752.
- _____. 1997. "The Effect of National Standards and Curriculum-Based Exams on Achievement." AMERICAN ECONOMIC REVIEW 87: 260-4.
- Blackburn, M. L., D. E. Bloom, and R. B. Freeman*. 1990. "The Declining Economic Position of Less Skilled American Males." In Gary Burtless, ed., A FUTURE OF LOUSY JOBS? THE CHANGING STRUCTURE OF U.S. WAGES, 31-67. Washington, D.C.: Brookings Institution.
- Cooper, Harris*. 1989. HOMEWORK. New York: Longman.
- Costrell, Robert M.* 1994. "A Simple Model of Educational Standards." AMERICAN ECONOMIC REVIEW 84: 956-71.
- Ehrenberg, Ronald G., Daniel D. Goldhaber, and Dominic J. Brewer*. 1995. "Do Teachers' Race, Gender and Ethnicity Matter? Evidence from the National Educational Longitudinal Study of 1988." INDUSTRIAL AND LABOR RELATIONS REVIEW 48: 547-61.
- Goldhaber, Daniel D., and Dominic J. Brewer*. 1997. "Why Don't Schools and Teachers Seem to Matter? Assessing the Impact of Unobservables on Educational Productivity." JOURNAL OF HUMAN RESOURCES 32: 505-23.
- Grogger, Jeff*. 1996. "Does School Quality Explain the Recent Black/White Wage Trend?" JOURNAL OF LABOR ECONOMICS 14: 231-53.
- Hanushek, Eric A.* 1996. "School Resources and Student Performance." In Gary Burtless, ed., DOES MONEY MATTER? THE EFFECT OF SCHOOL RESOURCES ON STUDENT ACHIEVEMENT AND ADULT SUCCESS, 43-73. Washington, D.C.: Brookings Institution.
- Hanushek, Eric A., et al.* 1994. MAKING SCHOOLS WORK: IMPROVING PERFORMANCE AND CONTROLLING COSTS. Washington, D.C.: Brookings Institution.
- Holmes, C. Thomas*. 1989. "Grade Level Retention Effects: A Meta-Analysis of Research Studies." In Lorrie A. Shepard and Mary Lee Smith, eds., FLUNKING GRADES: RESEARCH AND POLICIES ON RETENTION, 16-33. London: Falmer Press.
- Hoxby, Caroline M.* 1998. "What Do America's 'Traditional' Forms of School Choice Teach Us about School Choice Reforms?" Federal Reserve Bank of New York ECONOMIC POLICY REVIEW 4, no. 1.

REFERENCES (*Continued*)

- Lang, Kevin, and David Kropp.* 1986. "Human Capital Versus Sorting: The Effects of Compulsory Attendance Laws." *QUARTERLY JOURNAL OF ECONOMICS* 101: 609-24.
- Mincer, Jacob.* 1974. *SCHOOLING, EXPERIENCE, AND EARNINGS.* New York: Columbia University Press.
- Murnane, Richard J., Judith D. Singer, John B. Willett, James J. Kemple, and Randall J. Olsen.* 1991. *WHO WILL TEACH? POLICIES THAT MATTER.* Cambridge: Harvard University Press.
- National Center for Education Statistics.* 1991. *DIGEST OF EDUCATION STATISTICS, 1991.* Washington, D.C.: U.S. Government Printing Office.
- New York Times.* 1996. "Passing of Regents Exams to Be Required." April 24: A1, B9.
- Psacharopoulos, George.* 1985. "Returns to Education: A Further International Update and Implications." *JOURNAL OF HUMAN RESOURCES* 20: 583-604.
- Ravitch, Diane.* 1995. *NATIONAL STANDARDS IN AMERICAN EDUCATION: A CITIZEN'S GUIDE.* Washington, D.C.: Brookings Institution.
- Rouse, Cecilia Elena.* 1998. "Schools and Student Achievement: More Evidence from the Milwaukee Parental Choice Program." *Federal Reserve Bank of New York ECONOMIC POLICY REVIEW* 4, no. 1.
- Willis, Robert J., and Sherwin Rosen.* 1979. "Education and Self-Selection." *JOURNAL OF POLITICAL ECONOMY* 87: S7-S36.

The views expressed in this article are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System. The Federal Reserve Bank of New York provides no warranty, express or implied, as to the accuracy, timeliness, completeness, merchantability, or fitness for any particular purpose of any information contained in documents produced and provided by the Federal Reserve Bank of New York in any form or manner whatsoever.