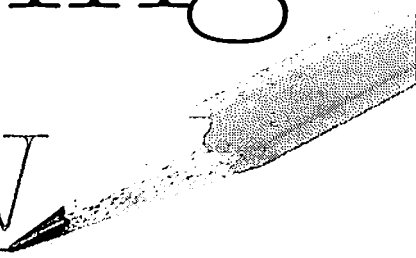
Creating a New Teaching Profession



EDITED BY
DAN GOLDHABER
AND JANE HANNAWAY

8

Teacher Deselection

Eric A. Hanushek

The national educational challenge was most forcefully articulated by the nation's governors in 1989. As they met in Charlottesville, Virginia, they felt the need of the nation to improve the performance of students—a need articulated a half decade previously in *A Nation at Risk* (National Commission on Excellence in Education 1983). And they declared that the United States should be first in the world in mathematics and science by the turn of the century (National Education Goals Panel 1991). The problem was that we had no experience to draw upon that would indicate how this could be done. In the intervening two decades we have come to recognize that improving teacher effectiveness is perhaps the only viable way to accomplish the governors' goals, but even there the policies and mechanisms are far from obvious.

This discussion provides a quantitative statement of one approach to achieving the governors' (and the nation's) goals—teacher deselection. Specifically, how much progress in student achievement could be accomplished by instituting a program of removing, or deselecting, the least-effective teachers? A variety of policies for hiring and retraining teachers have been proposed, but they have not been very successful in the aggregate, as student performance has not improved. At the same time, it is widely recognized that some teachers do a very poor job, and few people believe that the worst teachers can be transformed into good

teachers. What would happen if we simply adopted policies of systematically removing the most ineffective teachers?

Motivation

At the time of *A Nation at Risk*, the United States was not performing very well on international tests, but its school attainment far exceeded that in other countries. For example, 88 percent of U.S. students had finished high school, but only 72 percent of similarly aged students in Organisation for Economic Co-operation and Development (OECD) countries had done so.[1] Further, central features of the U.S. economic system—such as openness to trade, secure property rights and a well-developed legal system, and highly adaptable labor and product markets—insulated the economy from any flaws in the development of its labor force.

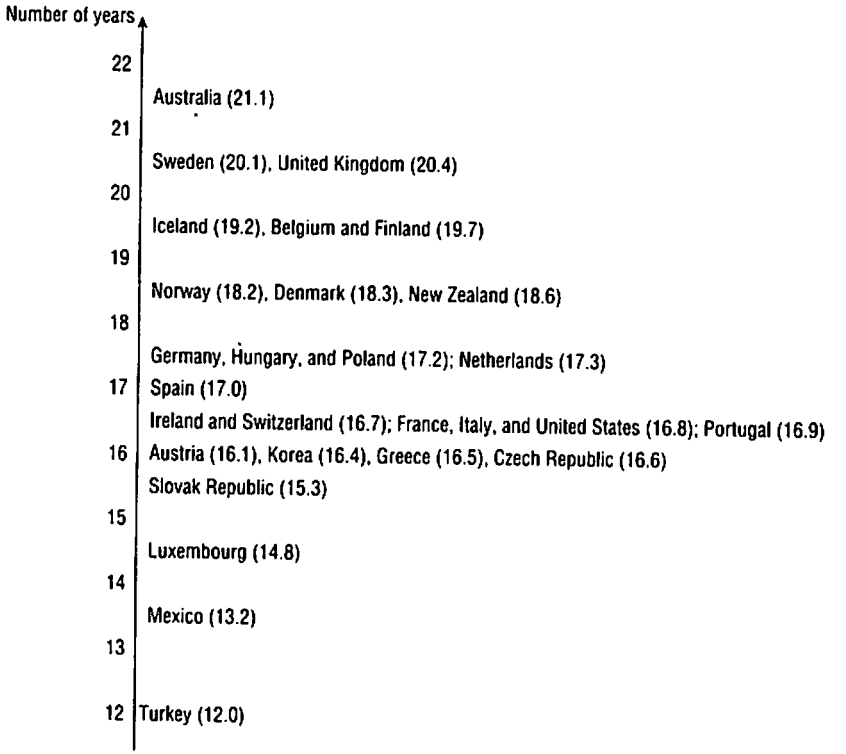
The world has changed dramatically since then. Other countries, intent on emulating the successes of the U.S. economy, have dramatically increased the school attainment of their populations. Figure 8.1 shows the expected school attainment in 2003. The United States falls noticeably below the average of 17.3 years for OECD countries, a remarkable shift in two decades. On other fronts, competitors have also been moving rapidly to improve their economic conditions to match those of the United States. As a result, other nations are currently much more competitive than at the time of *A Nation at Risk*—when the nation was told in unequivocal terms that the education system was not preparing our students to be competitive in the world.

The 1989 governors' meeting called for moving U.S. students up to the top of international rankings. But they did not attempt to describe what that would mean for the United States economy. We are now able to do that.

Start by considering how far behind the leader countries U.S. students are. Figure 8.2 shows the ranking of countries based on average mathematics score on the Program for International Student Assessment (PISA) tests in 2003.[2] U.S. students perform significantly below the OECD average. The top scorer on this assessment, Hong Kong, is two-thirds of a standard deviation (sd) ahead of the average U.S. student.

An improvement of 0.5 sd would move U.S. students close to the top—roughly where Canadian students fall and slightly behind such countries as Japan and the Netherlands.

Figure 8.1. Education Expectancy, 2003



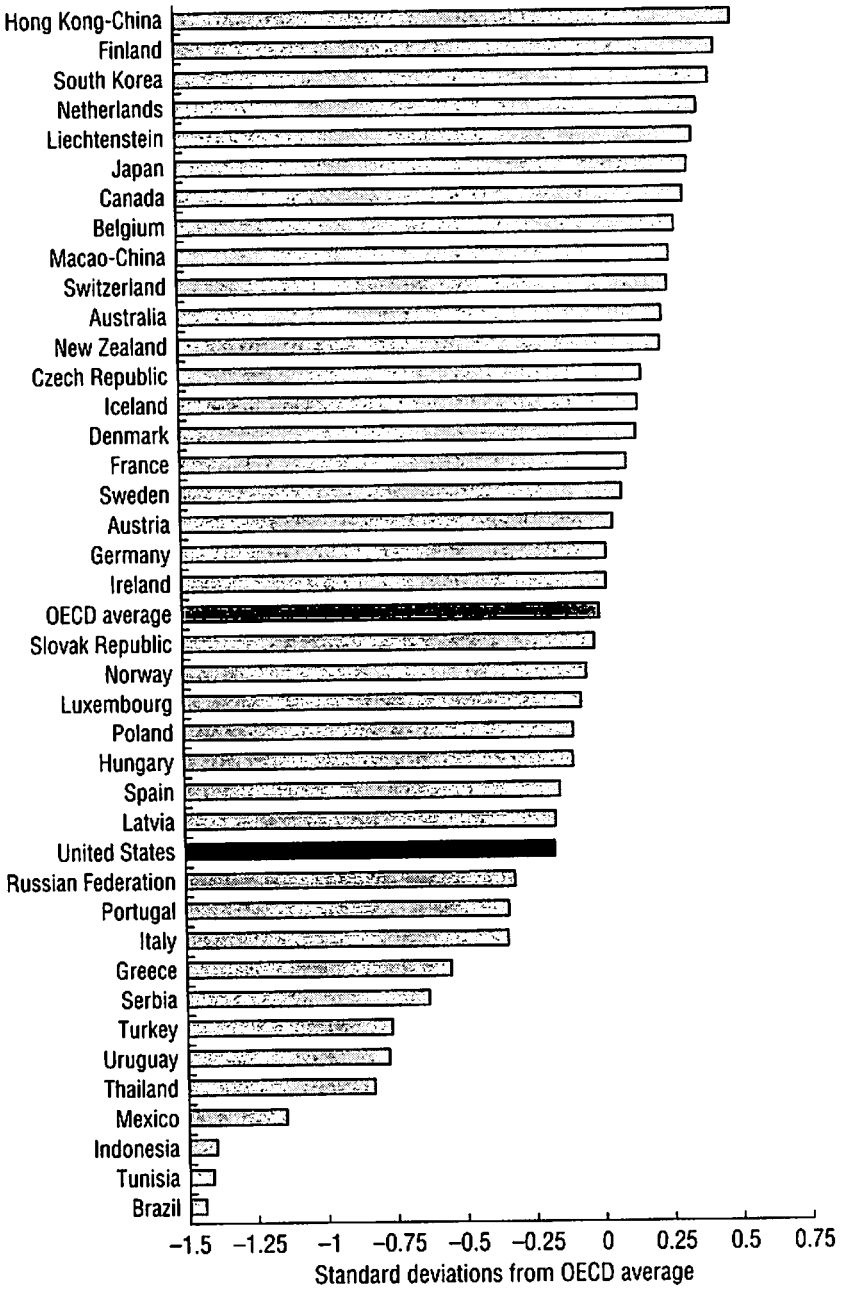
Source: OECD (2005), table C.1.1.

Note: Number of years includes all levels of education from primary education to adult life, under current conditions, excluding education for children under the age of 5.

Explanation: In Portugal, a 5-year-old child can expect to be enrolled during 16.9 years over his or her lifetime.

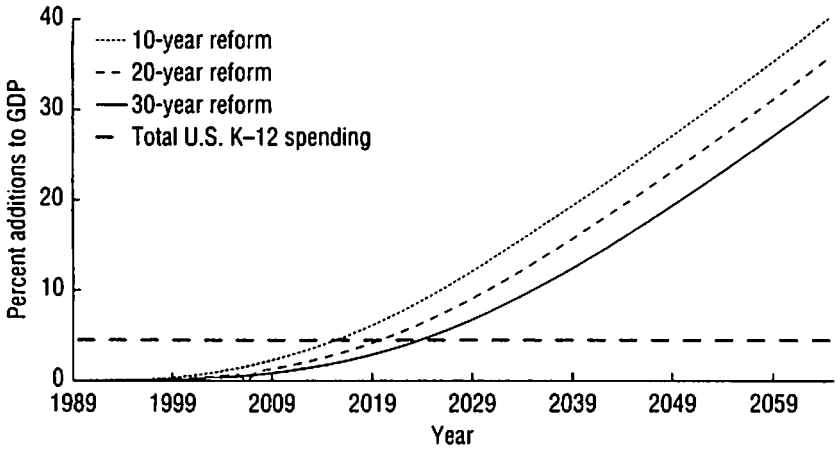
What would a 0.5 sd improvement mean for the U.S. economy? Recent analysis of how economic growth is affected by having a better-educated population shows that the implications would be dramatic (Hanushek et al. 2008). Figure 8.3, reproduced from that paper, shows how meeting the 1989 governors' pledge through improving student performance by 0.5 sd (making U.S. students perform like Canadians) would have affected current and future gross domestic product (GDP).

Figure 8.2. Mathematics Performance on PISA, 2003



Source: OECD (2004).

Figure 8.3. Increases in GDP from Improving Student Learning by 0.5 Standard Deviation



Source: Hanushek et al. (2008).

The governors' pledge of achieving dramatic improvements by 2000 is equivalent to the curve for a 10-year reform plan in figure 8.3.[3] This curve shows that by 2015 we could have expected GDP to be more than 4.5 percent higher than will be obtained without student improvement. (U.S. student performance according to the National Assessment of Educational Progress is essentially similar today to what it was at the time of the governors' summit meeting.) *This addition to GDP is equivalent to the proportionate expenditure for our 2008 national spending levels on K–12 education.* Or, seen differently, GDP in 2008 was more than $14 trillion; that makes the 2 percent increase in 2008 GDP predicted by the 10-year reform plan curve equivalent to about $300 billion.

If the reforms had been begun at the time of the governors' meeting but had stretched out for a longer period before they obtained their results, the improvements in GDP would take commensurately longer but would still have powerful implications for the U.S. economy. For example, a reform plan that took 30 years to bring students up to the level of Canada's students would cover K–12 expenditures with the added economic outcome by 2024.

The final motivating element for this chapter is that so little was accomplished by the policies that followed the governors' meeting or the

Child Left Behind Act in the sense that it describes a set of teacher policies that could lead to the results sought under the accountability systems.

The implication of the discussion is that improvements in student outcomes—if they occur—would be expected to have powerful effects. At the same time, doing more of the policies we have been pursuing is unlikely to lead to the performance improvements we seek.

The Power of Effective Teachers

A first question—given the previous section—is whether or not achieving such gains could be feasible with realistic reform strategies. Recent research suggests that it is feasible but that it will take redirection of efforts.

One explanation for past failure is simply that we have not directed sufficient attention to teacher quality and teacher effectiveness. By many accounts, the quality of teachers is the key element to improving student performance. But the research evidence suggests that many of the policies that have been pursued have not been very productive. Specifically, although the policies may have led to changes in measured aspects of teachers, they have not improved the quality of teachers as measured by student performance.[4]

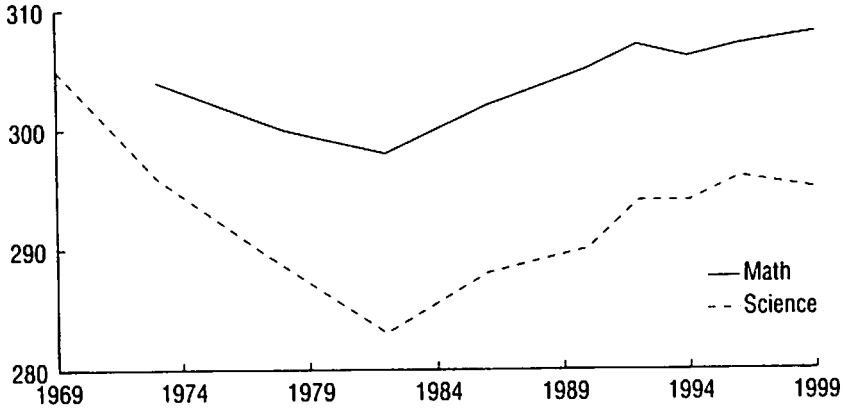
Rivkin, Hanushek, and Kain (2005) estimated differences in teacher quality on an output basis. Specifically, the concern is identifying good and bad teachers on the basis of their performance in obtaining gains in student achievement. An important element of that work is distinguishing the effects of teachers from the selection of schools by teachers and students, and the matching of teachers and students in the classroom. In particular, highly motivated parents search out schools that they think are good, and they attempt to place their children in classrooms where they think the teacher is particularly able (Hanushek, Kain, and Rivkin 2004a). Teachers follow a similar selection process (Hanushek, Kain, and Rivkin 2004b, 2004c). Thus, from an analytical viewpoint, it is difficult to sort out the quality of the teacher from the quality of the students in her classroom. The analysis of teacher performance in Rivkin and colleagues (2005) goes to great lengths to avoid contamination from any such selection and matching of kids and teachers.[5] In the end, that analysis estimates that the differences in annual achievement growth between an average and a good teacher in math are at least 0.11 sd of student achievement.[6]

previous call to action from *A Nation at Risk*. A variety of approaches have been pursued (Peterson 2003). These have involved expanding resources in many directions, including increasing real per-pupil spending by more than 50 percent since 1983. Yet U.S. performance has remained unchanged since 1970, when we started obtaining evidence from the National Assessment of Educational Progress (figure 8.4).

The aggregate picture is consistent with a variety of other studies indicating that resources alone have not yielded any systematic returns in student performance (Hanushek 2003). The character of reform efforts—at least until recently—can largely be described as "same operations with greater intensity." Thus, pupil–teacher ratios and class size have fallen dramatically, teacher experience has increased, and teacher graduate degrees have grown steadily, but these have not translated into higher student achievement. On top of these resources, a variety of programs have been introduced with limited aggregate success. The experience of the past several decades vividly illustrates the importance of true reform—that is, reform that actually improves student achievement.

The recent movement to a standards- and accountability-based reform may change this picture, but the evidence is yet to be clear. In any event, the discussion below is consistent with accountability and with the No

Figure 8.4. National Assessment of Educational Progress Scores for U.S. 17-Year-Olds, 1969–99



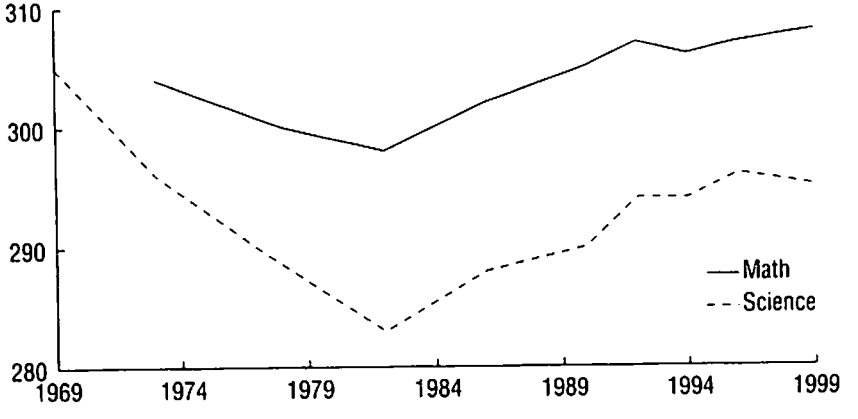
Source: National Center for Education Statistics, "The Nation's Report Card," http://nces.ed.gov/NATIONSREPORTCARD/.

previous call to action from *A Nation at Risk*. A variety of approaches have been pursued (Peterson 2003). These have involved expanding resources in many directions, including increasing real per-pupil spending by more than 50 percent since 1983. Yet U.S. performance has remained unchanged since 1970, when we started obtaining evidence from the National Assessment of Educational Progress (figure 8.4).

The aggregate picture is consistent with a variety of other studies indicating that resources alone have not yielded any systematic returns in student performance (Hanushek 2003). The character of reform efforts—at least until recently—can largely be described as "same operations with greater intensity." Thus, pupil–teacher ratios and class size have fallen dramatically, teacher experience has increased, and teacher graduate degrees have grown steadily, but these have not translated into higher student achievement. On top of these resources, a variety of programs have been introduced with limited aggregate success. The experience of the past several decades vividly illustrates the importance of true reform—that is, reform that actually improves student achievement.

The recent movement to a standards- and accountability-based reform may change this picture, but the evidence is yet to be clear. In any event, the discussion below is consistent with accountability and with the No

Figure 8.4. National Assessment of Educational Progress Scores for U.S. 17-Year-Olds, 1969–99



Source: National Center for Education Statistics, "The Nation's Report Card," http://nces.ed.gov/NATIONSREPORTCARD/.

Before going on, it is useful to put this lower-bound estimate of the variation in quality into perspective. By this quality estimate, if a student had a good teacher as opposed to an average teacher for four or five years in a row, the increased learning would be sufficient to close entirely the average gap between a typical low-income student receiving a free or reduced-price lunch and the average student who is not receiving free or reduced-price lunches.

Reasonable estimates of annual achievement growth (which are used throughout the following calculations) are actually higher than this lower bound—specifically, 0.20 to 0.30 sd. These larger estimates reflect likely differences in teacher quality among schools; the fact that the similarly conservative estimates for reading as opposed to math are 0.15 sd instead of 0.11 sd; and a series of other factors that bias the previously discussed estimate downwards (Rivkin et al. 2005).

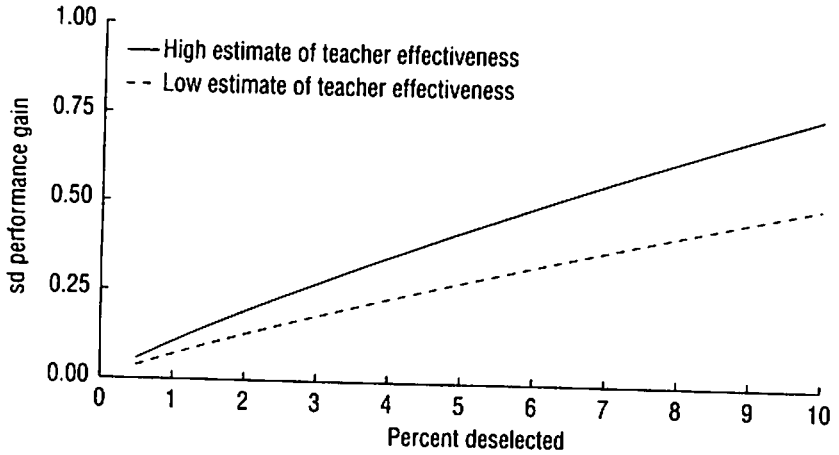
The implications of these differences are dramatic. Let us consider the impact of low-quality, or ineffective, teachers on student achievement. If the average learning growth each academic year is one grade-level equivalent, the estimates of variations in teacher quality indicate that the least effective 5 percent of teachers see gains that are at best two-thirds of a grade-level equivalent. The bottom 1 percent of teachers achieve no more than one-half of a grade-level equivalent in annual gains. (These calculations assume that 1.0 sd of teacher quality—moving from the center of the distribution to the 84th percentile—is 0.20 sd of student achievement; using a calculation of 0.30 sd makes these conclusions even more grim.)

External validation of these estimates comes from Hanushek (1992). The calculations of the low end of the distribution developed here are similar to the effects I previously calculated, but those estimates also suggest that the most conservative estimates may be too optimistic. The prior analysis of the range in performance in Gary, Indiana, schools suggests that the bottom 5 percent are no better than one-half grade-level equivalent in growth per academic year. These direct estimates of teacher differences are actually close to the higher estimates of teacher quality (0.30 sd of student achievement).

Clearly, the students with ineffective teachers are harmed. Students can probably recover from a single year of having a bottom 5 percent teacher, but a few years might lead to lasting problems—ones that dog students for a lifetime.

Let's look at the aggregate impact of the bottommost teachers. Figure 8.5 plots the impact on overall student learning of "deselecting"

Figure 8.5. Alternative Estimates of How Removing Ineffective Teachers Affects Student Achievement



Source: Author's calculations.

(i.e., moving out of the classroom) varying proportions of ineffective teachers. As an example, consider what would happen to average student performance if we could eliminate the least effective 5 percent of teachers from the distribution. The estimates of the impact of teachers on student achievement indicate that these students would on average gain 0.28–0.42 sd of performance.

These estimates of the importance of teacher quality permit some calculations of what would be required to yield the reforms discussed earlier. To begin with, consider what magnitude of teacher deselection might yield an improvement in student performance to the level of Canada (0.5 sd of student achievement). Figure 8.4 shows that eliminating the least effective 6–10 percent of teachers would bring student achievement up by 0.5 sd.

The estimates given here need to be put into the policy context. Consider a school with 30 typical teachers. These estimates suggest that eliminating the bottom two or three could boost student achievement up to the level of Canada's students.

This kind of policy is very consistent with the McKinsey evaluation of the policies found in high-performing school systems around the world (Barber and Mourshed 2007). Their evaluation suggests that the best school systems do not allow ineffective teachers to remain in the classroom

for long.[7] These conclusions are also consistent with more-local evidence, such as that for New York City, in Kane, Rockoff, and Staiger (2006) and the related policy prescriptions in Gordon, Kane, and Staiger (2006).

Policies of making active decisions on retention and tenure are, of course, alien to the current school system. A number of states currently have laws and regulations that lead to tenure decisions as early as two years into a teacher's career, with the mode being just three years (National Association of State Boards of Education 1997; National Council on Teacher Quality 2007). On top of that, the teacher evaluation process as typically seen is very cursory (Toch and Rothman 2008). Nonetheless, these are inconsistent with providing a quality education to all students, because some students must necessarily be relegated to these ineffective, and damaging, teachers.

The idea is that policies be put in place to identify the most-ineffective teachers and to move them out of the classroom. Developing such policies, negotiating them with teachers, and implementing them in the schools clearly take time. Thus, the reform paths in figure 8.3 sketch alternative time patterns for implementation that are likely to be relevant. Moreover, the prior calculations of impacts on student performance assume that all students have the higher-quality teachers for their entire K–12 school career. Thus, even if implemented immediately, it would take more than 10 years for any cohort to go through all schooling at the higher level. On the other hand, as noted earlier, had these policies been put in place as called for by the governors in 1989, we would today be reaping the unmistakable gains from the improved working of our national economy. Had we actually started with effective reform in 1983 as called for in *A Nation at Risk*, the gains to GDP would be sufficient to cover all of our expenditure on K–12 schooling.

It should also be noted that obtaining the gains from this policy presumes that the bottom 5 to 10 percent of the current distribution of teachers is eliminated *permanently*. To eliminate them permanently, it is necessary either to have a continuing deselection process or to upgrade the overall level of teacher effectiveness in the future. In particular, if hiring follows the current pattern, the new hires would have the same 5 to 10 percent of ineffective teachers, who would have to be deselected on an ongoing basis.

In the long run, it would probably be superior, however, to develop systems that upgrade the overall effectiveness of teachers. The difficulty is that past approaches have not proved very successful, at least as judged from stu-

dent outcomes. A variety of approaches have received widespread attention—induction programs, mentoring, professional development, and the like—and have been the focus of much policy interest. The interest in them has come substantially from the fact that they take current teachers and transform them into a more effective group. There has been little reason to believe that, at least as currently operated, these approaches are effective.[8] If they work, they could ensure that the pool of teachers is improved (and that continual deselection would not be needed). The key is "if they work."

The full impact of setting up a deselection process is difficult to project in the abstract. First, one might expect that a policy that selected and rewarded teachers on ability would alter who entered teaching. While more-risk-adverse people may shy away from teaching, a different group that wishes to be judged and evaluated by their contributions may come forward (see Hoxby and Leigh 2004). Moreover, there could be efficiency gains, for example, through improved professional development. (For more on the potential of using professional development to upgrade the quality of the teacher workforce, see chapter 11.) Today, when performance is not effectively evaluated, teachers and principals give little attention to the usefulness or quality of any professional development programs—and the result appears to be little average gains from the existing professional development. But, if classroom effectiveness mattered, teachers might be more engaged in selecting and participating in good professional development. Nonetheless, these ramifications are speculative, since we have little experience with how the market might operate with the introduction of true performance evaluation.

Cost Considerations

Policy initiatives on teacher deselection clearly alter the nature of the teaching contract. Today, few teachers are involuntarily separated from teaching, particularly after the probationary period. (In fact, some of the nation's worst performers—California and the District of Columbia—require decisions on tenure at the end of a two-year probationary period, making it very difficult to evaluate teachers.) As a result, the possibility of deselecting ineffective teachers increases the risk of employment as a teacher. Attempting to change the quality of the overall distribution may require increased compensation to attract

new and more-effective teachers who are also willing to take the added employment risk.

Analyzing reform policies directly by cost is not feasible because we know very little about the supply function for teacher quality. While there has been some work on the cost of hiring teachers with different characteristics (such as experience or advanced degrees), these characteristics do not readily translate into teacher quality (Hanushek and Rivkin 2004).

There are alternative ways to consider the costs of any policies aimed at improving the teaching force. Perhaps the simplest is to use the prior calculations of the benefits to provide an estimate of the upper bound on the feasible expenditure for new policies (based on the simple idea that costs must be less than benefits in order for the policy to be efficient).

Much of the current discussion of teacher quality is centered on statements about the overall level of salaries. It seems clear that teacher salaries have slipped relative to alternative earnings of college-educated workers, particularly for women (Hanushek and Rivkin 1997, 2004, 2006).[9] For various reasons, however, this does not give much policy guidance for the current discussions. In simplest terms, we do not know how teacher quality responds to different levels of salaries (Hanushek and Rivkin 2004). Moreover, policies that simply raised salaries across the board (even if advanced as a way to increase the attractiveness of the profession) would almost certainly slow any reform adjustments, because they would lower teacher turnover and make it more difficult to improve quality through new hiring.

Nonetheless, the aggregate growth numbers suggest that the annual growth dividend from an effective reform plan would cover most conceivable program costs over a relatively short period. Figure 8.3 shows an increase in GDP from improved student performance with varying implementation periods. Consider what real reform does. The "Canadian" improvement plan previously described that reached its goal in 20 years would already yield GDP that was 1.6 percent higher at the end of the 20-year implementation period. In other words, a 20-year reform plan begun in 2010 would yield this higher 1.6 percent level of GDP in 2030. In the U.S. economy in 2005, 1.6 percent amounted to $200 billion. That year, total spending on instructional salaries and benefits was just $233 billion (Snyder, Dillow, and Hoffman 2008, table 169). In other words, the increased GDP through improved student achievement would almost immediately cover current teacher salaries and benefits

fully—suggesting considerable room to pay for better teachers and to compensate for the higher risk of entering teaching.

The conclusion of the cost considerations is simple. The benefits from quality improvements are very large. Thus, they can support large and expansive incentive programs *if the programs work.* U.S. schools have, in fact, expanded in a variety of ways over the past four decades—real expenditures per pupil in 2000 are more than three times those in 1960. It is just that these past programs have not led to significant improvements in student performance. Put another way, the benefits do not justify all types of expenditure. They do justify many conceivable programs if they can be shown to be effective.

Conclusion

Many discussions of teacher policies concentrate almost entirely on upgrading the effectiveness of current teachers. This involves special induction programs, mentoring, ongoing professional development, and the like. Past approaches have not proved to be generally effective. Good programs have been difficult to implement on a broad scale, as seen through the results in student achievement.

This analysis points to the large cost of allowing the most-ineffective teachers to remain in the classroom. The bottom end of the teacher force is harming students. Allowing ineffective teachers to remain in the classroom is dragging down the nation.

If the bottom end of the teacher distribution cannot be improved through various remediation efforts, the alternative is more active deselection policies that trim off the least effective teachers. What stands out from an analysis of the impact of teachers on achievement is that relatively modest changes in the bottom end of the distribution have enormous implications for the nation.

It is unclear why we permit a small group of teachers to do such large damage. The majority of teachers are effective. They are able to compete with teachers virtually anywhere else in the world. Yet these effective teachers are lumped in with a small group of completely ineffective teachers, who are permitted to continue damaging students' educational experiences.

The problem does not seem to be with identifying these ineffective teachers. Some evidence shows that principals are able to identify these

bottom performers.[10] It is almost certainly true that there is even broader recognition of the bottom teachers—by principals, other teachers, and parents. Instead, it seems simply to be a lack of will to act on readily available information.

NOTES

1. These numbers reflect comparisons for the population born from 1961 to 1970; see OECD (2007), table A1.2a. These statistics, however, overstate the U.S. situation because they include GED completion in the calculation of high school success. This problem, along with other measurement issues, is discussed in Heckman and LaFontaine (forthcoming).

2. PISA is testing conducted by the OECD on international students for all OECD countries and a selection of other countries that voluntarily participate. PISA tests a random sample of 15-year-olds in each country. These tests, now on a three-year cycle, assess math, science, and reading skills. Alternative assessments at different ages are provided by TIMSS (Trends in International Mathematics and Science Study). The TIMSS shows some different comparative results, with U.S. 4th-graders doing relatively well, U.S. middle-school students in the middle, and U.S. 12th-graders in the bottom rankings.

3. For a description of how these calculations are accomplished, see Hanushek and Woessmann (2008, 2009).

4. For a review of existing literature, see Hanushek and Rivkin (2004, 2006). They describe various attempts to estimate the impact of teacher quality on student achievement.

5. To do this, they concentrate entirely on differences among teachers within a given school to avoid the potential impact of parental choice of schools. Moreover, they employ a strategy that compares grade-level performance across different cohorts of students, so the matching of students to specific teachers in a grade can be circumvented. As such, it is very much a lower-bound estimate on differences in teacher quality.

6. For this calculation, a teacher at the mean of the quality distribution is compared to a teacher 1.0 sd higher in the quality distribution (84th percentile), labeled a "good teacher." The estimate of 0.11 comes from the mathematics estimates. The comparable estimates for reading are 0.15 (revised estimates). These results relate specifically to grades 4–7. It is plausible that the distribution of teacher quality at high school is even more dispersed than shown here.

7. The method of ensuring good teachers, according to this study, does depend on the country. Some of the highest performing countries do this largely at entry by selecting from the very top of the pool of college graduates. Others have particularly effective professional development programs. For the United States, moving to selection of teachers from the top of the new graduate distribution appears infeasible. See Hanushek and Rivkin (2004).

8. Recent high-quality studies cast doubt on arguments about professional judgment and about various teacher induction and mentoring programs. See Garet et al. (2008) and Isenberg et al. (2009).

9. There is a current debate about how salaries of teachers compare to those in different professions; see Podgursky (2003) and Allegretto, Corcoran, and Mishel (2004).

10. See, for example, Armor et al. (1976), Jacob and Lefgren (2006), or Murnane (1975).

REFERENCES

Allegretto, Sylvia A., Sean P. Corcoran, and Lawrence Mishel. 2004. *How Does Teacher Pay Compare? Methodological Challenges and Answers.* Washington, DC: Economic Policy Institute.

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

Barber, Michael, and Mona Mourshed. 2007. *How the World's Best-Performing School Systems Come Out on Top.* McKinsey and Company.

Garet, Michael S., Stephanie Cronen, Marian Eaton, Anja Kurki, Meredith Ludwig, Wehmah Jones, Kazuaki Uekawa, et al. 2008. *The Impact of Two Professional Development Interventions on Early Reading Instruction and Achievement.* NCEE 2008-4030. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Gordon, Robert, Thomas J. Kane, and Douglas O. Staiger. 2006. "Identifying Effective Teachers Using Performance on the Job." Hamilton Project discussion paper. Washington, DC: The Brookings Institution.

Hanushek, Eric A. 1992. "The Trade-Off between Child Quantity and Quality." *Journal of Political Economy* 100(1): 84–117.

———. 2003. "The Importance of School Quality." In *Our Schools and Our Future: Are We Still at Risk?* edited by Paul E. Peterson (141–73). Stanford, CA: Hoover Institution Press.

Hanushek, Eric A., and Steven G. Rivkin. 1997. "Understanding the Twentieth-Century Growth in U.S. School Spending." *Journal of Human Resources* 32(1): 35–68.

———. 2004. "How to Improve the Supply of High-Quality Teachers." In *Brookings Papers on Education Policy 2004,* edited by Diane Ravitch (7–25). Washington, DC: Brookings Institution Press.

———. 2006. "Teacher Quality." In *Handbook of the Economics of Education,* edited by Eric Hanushek and Finis Welch (1051–78). Amsterdam: Elsevier.

Hanushek, Eric A., and Ludger Woessmann. 2008. "The Role of Cognitive Skills in Economic Development." *Journal of Economic Literature* 46(3): 607–68.

———. 2009. "Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation." Working Paper 14633. Cambridge, MA: National Bureau of Economic Research.

Hanushek, Eric A., John F. Kain, and Steve G. Rivkin. 2004a. "Disruption versus Tiebout Improvement: The Costs and Benefits of Switching Schools." *Journal of Public Economics* 88(9/10): 1721–46.

———. 2004b. "The Revolving Door." *Education Next* 4(1): 77–82.

———. 2004c. "Why Public Schools Lose Teachers." *Journal of Human Resources* 39(2): 326–54.

Hanushek, Eric A., Dean T. Jamison, Eliot A. Jamison, and Ludger Woessmann. 2008. "Education and Economic Growth: It's Not Just Going to School but Learning Something While There That Matters." *Education Next* 8(2): 62–70.

Heckman, James J., and Paul A. LaFontaine. Forthcoming. "The American High School Graduation Rate: Trends and Levels." *Review of Economics and Statistics.*

Hoxby, Caroline M., and Andrew Leigh. 2004. "Pulled Away or Pushed Out? Explaining the Decline of Teacher Aptitude in the United States." *American Economic Review* 94(2): 236–40.

Isenberg, Eric, Steven Glazerman, Martha Bleeker, Amy Johnson, Julieta Lugo-Gil, Mary Grider, Sarah Dolfin, and Edward Britton. 2009. *Impacts of Comprehensive Teacher Induction: Results From the Second Year of a Randomized Controlled Study.* NCEE 2009-4072. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Jacob, Brian A., and Lars Lefgren. 2006. "When Principals Rate Teachers." *Education Next* 6(2): 59–69.

Kane, Thomas J., Jonah E. Rockoff, and Douglas O. Staiger. 2006. "What Does Certification Tell Us about Teacher Effectiveness? Evidence from New York City." Working Paper 12155. Cambridge, MA: National Bureau of Economic Research.

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

National Association of State Boards of Education. 1997. "Teacher Tenure." *Policy Updates* 5(3): 1.

National Commission on Excellence in Education. 1983. *A Nation at Risk: The Imperative for Educational Reform.* Washington, DC: U.S. Government Printing Office.

National Council on Teacher Quality. 2007. *State Teacher Policy Yearbook, 2007.* Washington, DC: National Council on Teacher Quality.

National Education Goals Panel. 1991. *Building a Nation of Learners, 1991.* Washington, DC: National Education Goals Panel.

OECD. See Organisation for Economic Co-operation and Development.

Organisation for Economic Co-operation and Development (OECD). 2004. *Learning for Tomorrow's World: First Results from PISA 2003.* Paris: OECD.

———. 2005. *Education at a Glance: OECD Indicators 2005.* Paris: OECD.

———. 2007. *Education at a Glance: OECD Indicators 2007.* Paris: OECD.

Peterson, Paul E., ed. 2003. *Our Schools and Our Future: Are We Still at Risk?* Stanford, CA: Hoover Institution Press.

Podgursky, Michael J. 2003. "Fringe Benefits." *Education Next* 3(3): 71–76.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica* 73(2): 417–58.

Snyder, Thomas D., Sally A. Dillow, and Charlene M. Hoffman. 2008. *Digest of Education Statistics 2007.* NCES 2008-022. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

Toch, Thomas, and Robert Rothman. 2008. "Rush to Judgment: Teacher Evaluation in Public Education." Washington, DC: Education Sector.