CHAPTER 2

Boosting Teacher Effectiveness

Eric A. Hanushek

Over the last two decades, research on student achievement has pinpointed the central role of teachers. While other factors—families, peers, neighborhoods—are obviously elements in a student's learning, it is the school and particularly the teachers and administrators who are given the public responsibility for the education of our youth. There is a general consensus that improving the effectiveness of teachers is the key to lifting student achievement, although questions remain about how best to do this.

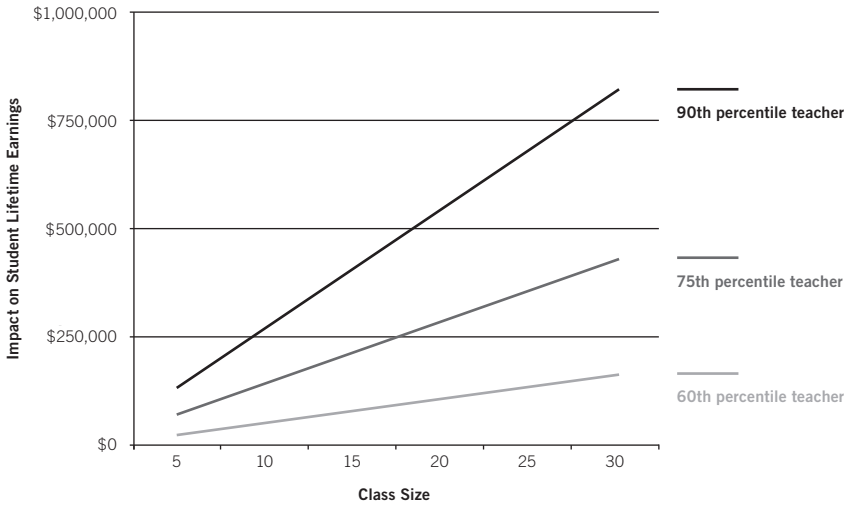
A key element in focusing attention on the importance of teacher effectiveness was research that took an outcomes-based perspective.[1] By looking at differences in the growth of student achievement across different teachers instead of concentrating on just the background and characteristics of teachers, it was possible to identify the true impact of teachers on students. This work, now generally called value-added analysis, demonstrated that some teachers consistently get greater learning gains year after year than other teachers. In fact, the average learning gains associated with a teacher provide a convenient metric for teacher effectiveness.

We now have a substantial number of studies that indicate clearly how much difference teacher effectiveness makes to student outcomes. In one study of mine, teachers near the top of the quality distribution got an entire year's worth of additional learning out of their students compared to those near the bottom.[2] That is, a good teacher will get a gain of 1.5 grade level equivalents while a bad teacher will get 0.5 year during a single academic year. Importantly, this analysis considered kids just from minority and poor inner-city families, indicating that family background is not fate and that good teachers can overcome deficits that might come from poorer learning conditions in the home.

A second perspective comes from combining existing quantitative estimates of how differences in teacher quality relate to achievement gaps by race or income.[3] Moving from an average teacher to one at the eighty-fourth percentile of teacher quality (i.e., moving up one standard deviation in teacher quality) would close somewhere between one-quarter and one-third of the average gap in math achievement between kids eligible for free and reduced-price lunches and those with higher incomes. Said differently, having a good teacher as opposed to an average teacher for three to four years in a row would, by available estimates, close the achievement gap by income. Closing the black-white achievement gap, which is a little larger than the average income gap, would take good teachers three and a half to five years in a row.

Perhaps the most valuable way to see differences for the subsequent discussion of salaries is to calculate the impacts of effective teachers on the future earnings of students.[4] A teacher who raises the achievement of a student will tend, other things being equal, to raise earnings throughout that student's work life. We can in fact calculate the economic impact on the student from analyses of how achievement translates into higher incomes. Using 2010 earnings, for example, a teacher in the seventy-fifth percentile would on average raise each student's lifetime income by somewhat

FIGURE 1. Impact on Student Lifetime Incomes by Class Size and Teacher Effectiveness (compared to average teacher)



more than $14,300 when compared to the average teacher. (All calculations are based on present values at the time of high school graduation where future incomes are discounted at 3 percent per year). But, this is not fully what the seventy-fifth percentile teacher contributes, because each student in the class can expect the same enhanced income. Thus, with a class of twenty-five students, this teacher would add $358,000 in future income compared to an average teacher.

Figure 1 shows the total contribution of teachers at the sixtieth, seventy-fifth, and ninetieth percentile of teacher effectiveness with varying class sizes. Excellent teachers add over $800,000 to the future incomes of students in a class of thirty. Even a teacher just above average at the sixtieth percentile would add over $100,000 to a class of twenty students.

These are calculations for each school year. Each and every year throughout their careers that these above-average teachers are

teaching adds hundreds of thousands of dollars to their students' future incomes. They also parallel a recent set of direct estimates of income effects that comes from linking teacher "value-added" to income tax records.[5]

But, there is the darker side. Below-average teachers are subtracting from student earnings at a similar rate. The tenth percentile teacher, compared to an average teacher, subtracts over a half million dollars per year for each twenty students he or she teaches. For the tenth, twenty-fifth, and fortieth percentile teacher, one simply has to put a minus sign in front of the values seen in figure 1.

From these different perspectives, the answer is the same: teachers have an enormous influence on students and on their futures.

Of course, there are two ways to look at the policy relevance of these figures. One is to assume that the current stock of teachers is fixed so that it is all just a zero-sum game—what one student gains, another student loses by being stuck with a below-average teacher. In such a case teacher policy would amount to deciding one way or another who gets the good teachers and who gets the bad teachers. It could, for example, be decided by market forces that allocate teachers to schools, or it could be decided by regulatory approaches, perhaps emanating from court cases and the like.

When viewed as just a distributional issue, the country as a whole would be no better off in terms of overall productivity by potential policy changes (even if the outcomes were viewed as more just). This result is not consistent with the primary concerns about education that relate to the overall productivity and output of the nation.

The second way to look at the prior calculations is more consistent with our investment notions about education. The impact of teachers on lifetime earnings is meant to signal how the productivity of individuals changes with different skills. The figures indicate the gains that would accrue to having more teachers of the type

found in the top of the current effectiveness distribution. Similarly, they indicate the gains that would accrue to having fewer teachers near the bottom of the current distribution. In other words, a policy that increased the average quality of the teacher distribution from that currently in place would yield potentially large overall gains to the economy (and potentially improve distributional matters at the same time).

To illustrate this, I have used the information available about the varying effectiveness of the teacher force to understand the aggregate impact of policies aimed at eliminating the worst teachers (an issue with direct policy implications as discussed below). I have projected the achievement impact of replacing varying percentages of the bottom teachers with average teachers.[6] By eliminating just the bottom 5–8 percent of teachers, the available research suggests that US achievement could climb to the level of Canadian achievement (as measured by international assessments of math and science).

Paul E. Peterson, Ludger Woessmann, and I have developed the economic implications of improving achievement to the level of Canada.[7] This analysis is based on the strong impacts of worker skills on future economic growth. It suggests that all workers in the United States could, by historical results, expect an average increase in their paychecks of 20 percent for each of the next eighty years. An alternative way to look at this is that the current fiscal problems could be readily solved by improved education that led to improved growth. Clearly, similar to the individual findings, there are substantial economic gains that seem apparent from policies that upgrade the quality of teachers.

While less well-developed, an increasing body of evidence points to the importance of principal quality.[8] The currently available research, based on the value added by a principal to achievement of students in the school, indicates that the principal may have an impact on achievement similar to that of teachers, although the principal affects the entire school.

Current Policy Discussions

Policy debates have changed swiftly to incorporate the research evidence on teachers.[9] It is difficult to enter into any school policy discussion that does not touch on the issue of teacher quality.

Moreover, the character of the discussions has become much more sophisticated and knowledgeable. The naïve calls for "highly qualified teachers" in the No Child Left Behind act have been replaced by recognition that credentials and qualifications—the objects of past policies—are not closely related to teacher effectiveness in the classroom. While there has been no rush to eliminate salary differentials based on advanced degrees (about 10 percent of all teacher salary payments), they have become a greater part of the discussion.

Similarly, a teacher's classroom experience after the first few years has been shown to have no effect on teacher performance. There has been little discussion of eliminating the longevity portion of teachers' salaries, even though over one-quarter of the total wage bill goes to bonuses for teachers with greater than two years of experience (around the cutoff in the evidence about the returns to experience). But there has been intense discussion of LIFO provisions—last in, first out—in laws and contracts that govern separations during force reductions. These policies are closely related to the evidence on effectiveness and experience. The use of LIFO rules instead of ones based on teacher effectiveness have been shown to increase the number of teachers who must be dismissed and to dramatically alter the quality of dismissals when compared to policies based on effectiveness.[10]

While considerable discussion exists on how we might want to change schools of education, little of this is directly related to the performance of students. Indeed, we have just rudimentary evidence on whether some schools of education do a better job than others. There is suggestive information in the fact that there is not

very much difference in average effectiveness by teachers' routes into their careers (certified vs. non-certified).[11]

Similarly, many would like to use improved professional development to upgrade the teaching force, but many questions about the efficacy of this remain.[12] Further (scientific) research on the issues surrounding professional development could prove helpful in deciding the overall thrust of teacher improvement policies.

Importantly, with the recognition of the importance of teacher quality has come a new interest in how labor laws and teacher contracts affect student outcomes.[13] The turmoil in Wisconsin got the most attention as the state limited bargaining to just wages and benefits and removed larger issues such as class sizes and teacher assignment policies. Partly because of Wisconsin and partly on their own, a number of other states entered into active discussions of state restrictions on teaching.

A central part of much of the teacher quality discussion has been the use of value-added measures of quality. The value-added measures are designed to provide estimates of the independent effect of the teacher on the growth in a student's learning and to separate this from other influences on achievement such as families, peers, and neighborhoods. The validity and reliability of these measures have been widely debated and are the subject of considerable current research.[14]

The discussions range across a number of statistical and policy issues. But the discussion was accelerated when the *Los Angeles Times* and the *New York Times* (among others) published the names and value-added rankings of thousands of local teachers. The public attention to variations in teacher effectiveness led to an uproar—an uproar that helped focus the policy discussion and local bargaining.

Attention to test scores in the value-added estimation raises issues of the narrowness of the tests, of the limited numbers of teachers in tested subjects and grades, of the accuracy of linking

teachers and students, and of the measurement errors in the achievement tests. Each is an important issue that has fueled continuing research efforts. This subsequent research is helping to define how best to use the statistical evidence on teacher quality.

The value-added discussions have also opened new consideration of alternative ways of valuing teachers. While teachers have always been evaluated in some manner, it is clear that until recently the evaluations provided little information, particularly for making any personnel decisions.[15] Thus, efforts have been made to develop and use observational protocols that more accurately indicate classroom effectiveness.[16]

A closely related discussion has revolved around the use of performance pay. Teachers are currently paid according to experience and to possession of an advanced degree, neither of which is closely related to classroom effectiveness. The argument has long been made that at least a portion of pay should reflect merit in order to provide incentives for teachers to do better. This idea led to a somewhat ill-conceived experiment by Vanderbilt researchers in which a randomly selected group of teachers received bonuses based on the performance of their students.[17] When compared to students of teachers not receiving any bonuses, students of those with the possibility of performance pay did no better. This study demonstrated that offering a bonus for better performance to existing teachers has very little influence on what they do. This is exactly what has been shown by the multiple studies of merit pay that focus on the impact of relatively small bonuses for current teachers on their performance in the classroom. The simplest interpretation is that almost all current teachers are indeed working to do the best that they can.

At the same time, this is not a demonstration that salaries have no effect. Both the level of salaries and the pattern of salaries across teachers affect who enters and who stays in teaching. Higher salaries and a greater relationship to performance would attract a different group of people into teaching. Indeed, the impact of salaries

on selection into teaching is the key issue for those who think that performance pay is important.

Nonetheless, the availability of this "gold standard" study has allowed the unions and the schools to argue that performance pay has been tried and simply has not worked. This situation actually demonstrates a further issue in making evidenced-based policy. It is often possible to find or to interpret evidence in order to support very different positions. This in fact has made moving to rational policy positions more difficult, particularly in areas of personnel policy where vested interests are especially important.

The possibility of evidence being hijacked for the use of special interest groups serves to reinforce the need for continued research and evaluation. Only superior and more reliable evidence can top the biased use of evidence.

The Prospects for Further Improvement

The world of education is moving steadily toward reliance on evidence, even with the possibility for misinterpretation. Moreover, the evidence on teacher quality issues is beginning to win the day.

The movement toward better overall policy is seen directly in state actions. For example, all states except California had unique student identifiers in 2011; thirty-five had unique teacher identifiers that allowed linking teachers to students.[18] Between 2009 and 2011, twenty-six states moved to include evidence of student learning in teacher evaluations, and ten states mandated that student learning would be the preponderant criterion in local evaluations.

In teacher tenure decisions, there has been considerable recent progress. More and more states are moving to require evidence of teacher effectiveness and to extend the minimum number of years for tenure. About a third of states also support differential pay in shortage subject areas and do not have regulatory language blocking differential pay. Similarly, about a third of states support differentially rewarding effective teachers.

While there is a considerable way to go in expanding and refining these changes, the pattern of state policies toward effective teachers has changed dramatically in recent years.

And there is a new sense of forward movement at the local level. Perhaps the best story comes from Washington, DC. This district, by far the worst in the nation, went through agonizing battles between Michelle Rhee (the former chancellor of Washington public schools) and the unions. But four years ago the unions agreed to a new contract that introduced both value-added and observational evaluations and that used them in personnel decisions. At this time about one thousand teachers have received substantial increases in their base salaries because of continued top performance. But close to 500 teachers have been dismissed because of continued poor performance. The whole evaluation system is continually being developed and improved, but it has reached a level of acceptance that bodes well for the future.

Importantly, there is now direct evidence that the Washington, DC, personnel policies are paying off. Thomas Dee and James Wyckoff found that dismissal threats increased the voluntary attrition of low-performing teachers by more than 50 percent.[19] Additionally, low-performing teachers who stayed improved their performance significantly, as did high-performing teachers who were in the range to get bonuses.

Similarly, the Los Angeles Unified School District has moved to remove around one hundred poorly performing teachers. While this remains small compared to the total number of teachers in Los Angeles, it is orders of magnitude larger than what was seen just a couple of years ago.

Many states and localities are developing what must be thought of as experimental programs for ensuring teacher quality. The key to the future is validating and replicating the ones that prove successful and eliminating the ones that do not. Doing this requires a strong research and evaluation activity to match the policy experimentation.

Notes

1. Eric A. Hanushek, "Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro-Data," *American Economic Review* 60, no. 2 (May 1971): 280–288.
2. Eric A. Hanushek, "The Trade-Off Between Child Quantity and Quality," *Journal of Political Economy* 100, no. 1 (February 1992): 84–117.
3. Eric A. Hanushek and Steven G. Rivkin, "The Distribution of Teacher Quality and Implications for Policy," *Annual Review of Economics* 4 (2012): 131–157.
4. Eric A. Hanushek, "Valuing Teachers: How Much is a Good Teacher Worth?" *Education Next* 11, no. 3 (Summer 2011).
5. Raj Chetty, John N. Friedman, and Jonah E. Rockoff, "The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood" (NBER working paper 17699, Cambridge, MA: National Bureau of Economic Research, December 2011).
6. Eric A. Hanushek, "The Economic Value of Higher Teacher Quality," *Economics of Education Review* 30, no. 3 (June 2011): 466–479.
7. Eric A. Hanushek, Paul E. Peterson, and Ludger Woessmann, *Endangering Prosperity: A Global View of the American School* (Washington, DC: Brookings Institution Press, 2013).
8. Gregory F. Branch, Eric A. Hanushek, and Steven G. Rivkin, "School Leaders Matter: Measuring the impact of Effective Principals," *Education Next* 13, no. 1 (Winter 2013): 62–69.
9. For example, John E. Chubb, *The Best Teachers in the World: Why We Don't Have Them and How We Could* (Stanford, CA: Hoover Institution Press, 2012).
10. For example, Donald Boyd, Hamilton Lankford, Susanna Loeb, and James Wyckoff, "Teacher Layoffs: An Empirical Illustration of Seniority versus Measures of Effectiveness," *Education Finance and Policy* 6, no. 3 (Summer 2011): 439–454.
11. For example, Thomas J. Kane, Jonah E. Rockoff, and Douglas O. Staiger, "What Does Certification Tell Us About Teacher

Effectiveness? Evidence from New York City," *Economics of Education Review* 27, no. 6 (December 2008): 615–631; and Donald J. Boyd, Pamela L. Grossman, Hamilton Lankford, Susanna Loeb, and James Wyckoff, "Teacher Preparation and Student Achievement," *Educational Evaluation and Policy Analysis* 31, no. 4 (December 2009): 416–440.

12. For example, Michael S. Garet, Stephanie Cronen, Marian Eaton, Anja Kurki, Meredith Ludwig, Wehmah Jones, Kazuaki Uekawa, Audrey Falk, Howard S. Bloom, Fred Doolittle, Pei Zhu, and Laura Sztejnberg, "The Impact of Two Professional Development Interventions on Early Reading Instruction and Achievement," National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences (Washington, DC: US Department of Education, September 2009); and Michael S. Garet, Andrew J. Wayne, Fran Stancavage, James Taylor, Marian Eaton, Kirk Walters, Mengli Song, Seth Brown, Steven Hurlburt, Pei Zhu, Susan Sepanik, and Fred Doolittle, "Middle School Mathematics Professional Development Impact Study: Findings After the Second Year of Implementation," NCEE 2011-4024 (Washington, DC: Institute of Education Sciences, April 2011).
13. Terry M. Moe, *Special Interest: Teachers Unions and America's Public Schools* (Washington, DC: Brookings Institution Press, 2011).
14. Steven Glazerman, Susanna Loeb, Dan Goldhaber, Douglas Staiger, Stephen Raudenbush, and Grover Whitehurst, "Evaluating Teachers: The Important Role of Value-Added," The Brookings Brown Center Task Group on Teacher Quality (Washington, DC: Brookings Institution Press, November 17, 2010).
15. For example, Daniel Weisberg, Susan Sexton, Jennifer Mulhern, and David Keeling, "The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness," 2nd ed. (New York: The New Teacher Project, 2009).

16. Thomas J. Kane, Daniel F. McCaffrey, Trey Miller, and Douglas O. Staiger, "Have We Identified Effective Teachers? Validating Measures of Effective Teaching Using Random Assignment," Measures of Effective Teaching project, Bill and Melinda Gates Foundation (January 2013).
17. Matthew G. Springer, Dale Ballou, Laura Hamilton, Vi-Nhuan Le, J.R. Lockwood, Daniel F. McCaffrey, Matthew Pepper, and Brian M. Stecher, "Teacher Pay for Performance: Experimental Evidence from the Project on Incentives in Teaching" (Nashville, TN: National Center on Performance Incentives, Vanderbilt University, 2010).
18. National Council on Teacher Quality, "State Teacher Policy Yearbook, 2011" (Washington, DC: National Council on Teacher Quality, 2012).
19. Thomas Dee and James Wyckoff, "Incentives, Selection, and Teacher Performance: Evidence from IMPACT" (NBER Working paper WP19529, Cambridge, MA: National Bureau of Economic Research, October 2013).