Caroline M. Hoxby, Harvard University,
2003 Distinguished Guest Lecturer.

Southern Economic Journal 2004, 71(2), 209–231

Distinguished Guest Lecture

Productivity in Education: The Quintessential Upstream Industry

Caroline M. Hoxby*

Using consistent test score data from the National Assessment of Educational Progress and data on per-pupil spending, I show that the productivity of American public schools fell by approximately half from 1970 to 2000. The most reliable international data also suggest that productivity in American public schools is lower than that of numerous other industrialized countries, including the remaining English-speaking ones. I explore explanations for the decline in productivity, including changing sociodemographics, Baumol's "cost disease," rising wages of female college graduates, the increasing emphasis on educating disadvantaged children, rising market power, and the education sector's relative decrease in pay for performance. I review evidence that suggests that schools raise their productivity and use of pay for performance when they face competition. I also describe results that indicate that individual teachers have important, distinctive effects on achievement.

JEL Classification: H400, H 700, I200, J240

1. Education as an Upstream Industry

Hardly a day goes that we do not hear about the way today's economy relies on human skills. We may be justifiably dubious about hyperbolic statements: Not everyone in America works in the "New Economy," a knowledge-based service-sector industry, such as software design. Nevertheless, America is, with every passing year, increasingly reliant on industries that use workers who have formal education, training, and the skills that allow a person to adapt rapidly to changing demands. Manufacturing has become substantially more oriented toward specialized production that produces customized products. There are many advantages of having an economy dominated by skill-oriented industries: They generate high per-capita incomes, and they tend to be environmentally "friendly." An upstream industry is any industry that produces inputs for other industries that are closer to the product market. Yet, when the words "upstream industry" appear, steel or petrochemicals often spring

* Harvard University, Department of Economics, Littauer Center 222, Cambridge, MA 02138, USA; E-mail choxby@kuznets.fas.harvard.edu.

Caroline M. Hoxby presented the Distinguished Guest Lecture at the 2003 Annual Meeting of the Southern Economic Association in San Antonio, Texas. She is a professor of economics at Harvard University and the director of the Economics of Education Program for the National Bureau of Economic Research. Hoxby has an undergraduate degree, a master's degree, and a doctorate in economics. She earned her master's degree in 1990 from the University of Oxford, which she attended on a Rhodes Scholarship, and her doctorate in 1994 from the Massachusetts Institute of Technology. Hoxby's research has received numerous awards, including a Carnegie Fellowship, a John M. Olin Fellowship, a National Tax Association Award, and a major grant from the National Institute of Child Health and Development.

to mind. However, the sector that produces education is probably the upstream industry on which skill-oriented economies most rely.

Many economic models link education to a country's growth. Nevertheless, thinking about the education sector as an upstream industry generates a new perspective on the education–growth relationship. We know that downstream industries locate where there is a relative abundance of the inputs on which they disproportionately rely. If the United States is to have future growth based on skilled industries, it must maintain its relative abundance of human capital.

Human capital is made, not naturally endowed, so in the long run, it will be relatively abundant where the education sector has high productivity—that is, where schools efficiently transform inputs into skill. For years, America did have a high-productivity elementary and secondary education sector: Compared to other countries, America spent only moderately on education, and yet its population completed an unusually large number of years of education. For more than thirty-five years, however, the extensive margin (more years of elementary and secondary education per person) has been exhausted in the United States. Moreover, on the extensive margin, all of the developed world has caught up or more than caught up to the United States (Organization for Economic Cooperation and Development [OECD] 2003). Achievement per dollar spent has thus become the key measure of productivity in the primary and secondary (K–12) education sector.

2. Measuring the Productivity of Public Primary and Secondary Schools

Measuring productivity in education is somewhat difficult. Indeed, measuring productivity in any industry is somewhat difficult. Education is so important, however, that we must try seriously to measure it. I will start with time-series data from the United States, which can give us the range of productivity that the sector has displayed in the last few decades. Then I will compare the United States with OECD and other countries in a recent year. This comparison will give us the range of productivity that exists in competing countries.

Even with all of the bells and whistles that can be included, this will not be an empirical exercise with the exactitude beloved by today's applied microeconomists. It is, however, sufficiently accurate to motivate the remaining discussion.

One measures the productivity of the education industry by dividing its outputs by its inputs. The inputs are relatively easy: Per-pupil spending is the best available measure. The outputs we can measure are limited, especially if we want to look at schools' productivity now. Test scores are available relatively quickly; lifetime earnings are not. Also, it has proven very difficult for researchers to identify how primary and secondary educational inputs contribute to later earnings, partly because of self-selection into further education and partly because many factors other than education affect later earnings. Thus, I will use test scores rather than earnings as the measure of output. Fortunately, the test scores of the sort I use have been shown to be strong predictors of later wages, employment, and other outcomes.[1] I will use test scores that are representative of the whole population of potential

[1] The National Assessment of Educational Progress is written by Educational Testing Services, which also writes the SATI, the SATII achievement tests, the Advanced Placement tests, and the achievement tests used in the National Education Longitudinal Study (NELS). A variety of studies have demonstrated that individuals' scores on these tests are correlated with their later earnings. For instance, even though the students in the NELS were only age 26 in 2000, their earnings already had a correlation of about 0.2 with their scores on the mathematics test they took in the 12th grade (author's calculations based on U.S. Department of Education 2002).

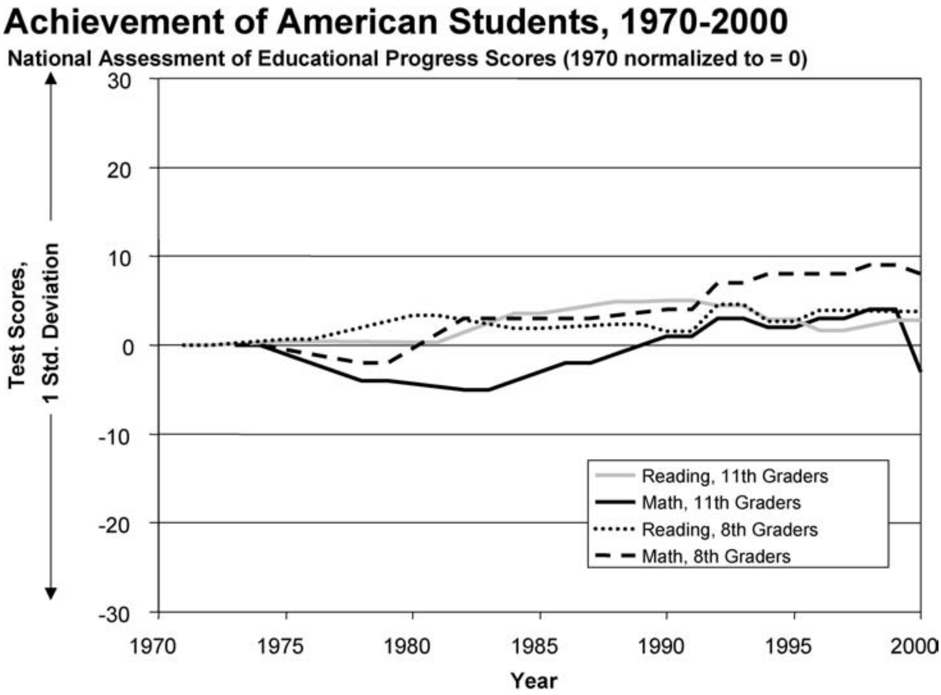

Figure 1. Achievement of American Students, 1970–2000 (National Assessment of Educational Progress Scores with 1970 normalized to zero). Source: Author's calculations using U.S. Department of Education (2003e) data.

students, not merely self-selected students who have chosen to remain in school past some common stopping point.

Let us first consider the United States, where I can make solid computations for the last three decades. Since 1970, American students have taken a series of tests called the National Assessment of Educational Progress (NAEP) that are specifically designed to track achievement from year to year. The NAEP is administered to very large samples of American students in a few grades.[2]

Figure 1 shows performance on the NAEP from 1971 to 2000 (U.S. Department of Education 2003e). I have normalized the 1970 scores to zero. For ease of interpretation, the height of the vertical axis is one standard deviation. That is, a student's score would have to rise from the very bottom of the figure to the very top of the figure if he were to move one standard deviation on the test. Clearly, regardless of the subject or grade we examine, performance has been flat.

One might worry that the flat test scores are the result of better schools working on students with worse sociodemographics. The data suggest that this worry is unnecessary. If one weights a student's scores by the number of students who were in his family background cell in 1970, one gets the line labeled "Using 1970 Socio-Demographics" in Figure 2.[3] If one weights a student's scores by the number of students who were in his family background cell in 2000, one gets the line labeled "Using 2000 Socio-Demographics" in Figure 2. (Notice that I have normed achievement so that actual achievement in 1970 is always zero.) If you have eagle eyes, you may be able to see that year 2000 sociodemographics are slightly more favorable than 1970 sociodemographics. This is largely a result

---

[2] The sample size in the NAEP ranges from about 15,000 to about 75,000 students, depending on the year.

[3] The background cells are based on race, Hispanic ethnicity, urbanicity, parents' income, parents' education, and foreign birth. These variables are available in the microdata for the National Assessment of Educational Progress.

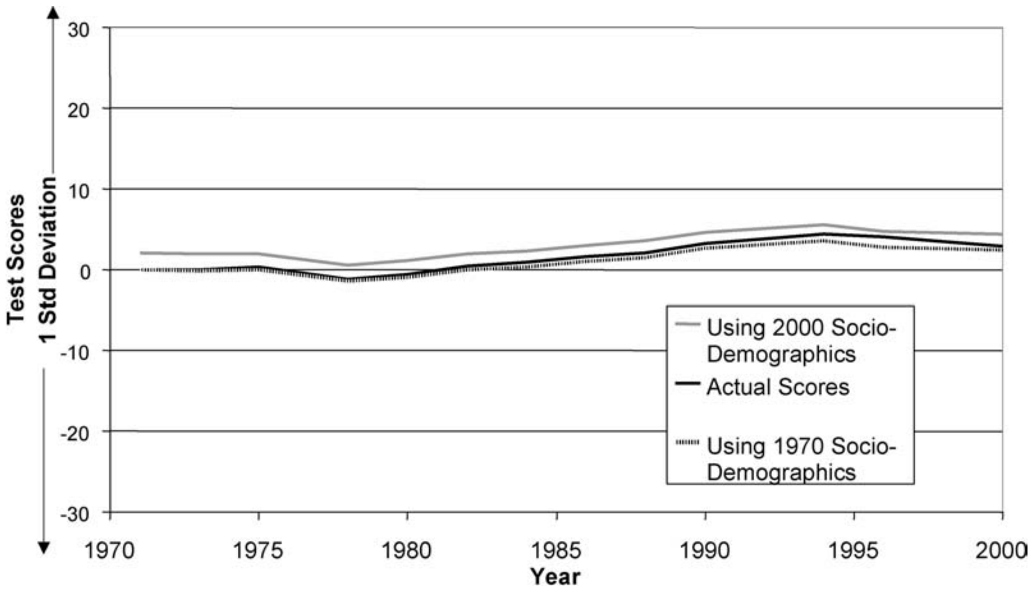

Figure 2. Achievement, holding students' sociodemographics constant, 1970–2000 (National Assessment of Educational Progress Scores, 1970 actual scores normalized to zero). Source: Author's calculations using U.S. Department of Education (2003e) data.

of the growth of family income in the United States. Nevertheless, the message of Figure 2 is that sociodemographics cannot easily be made to account for flatness of achievement.

Achievement may have been quite flat, but the same cannot be said for inputs into public primary and secondary education in the United States. Figure 3 shows per-pupil spending in public schools

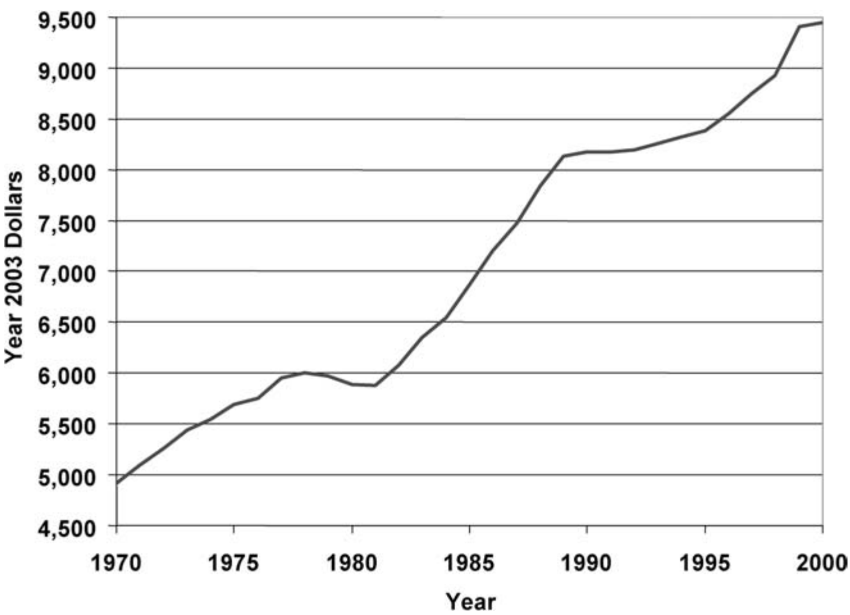

Figure 3. Average per-pupil spending in American public schools (year 2003 dollars, cost index is CPI). Source: Author's calculations using U.S. Department of Education (2004) and U.S. Department of Labor (2004) data.

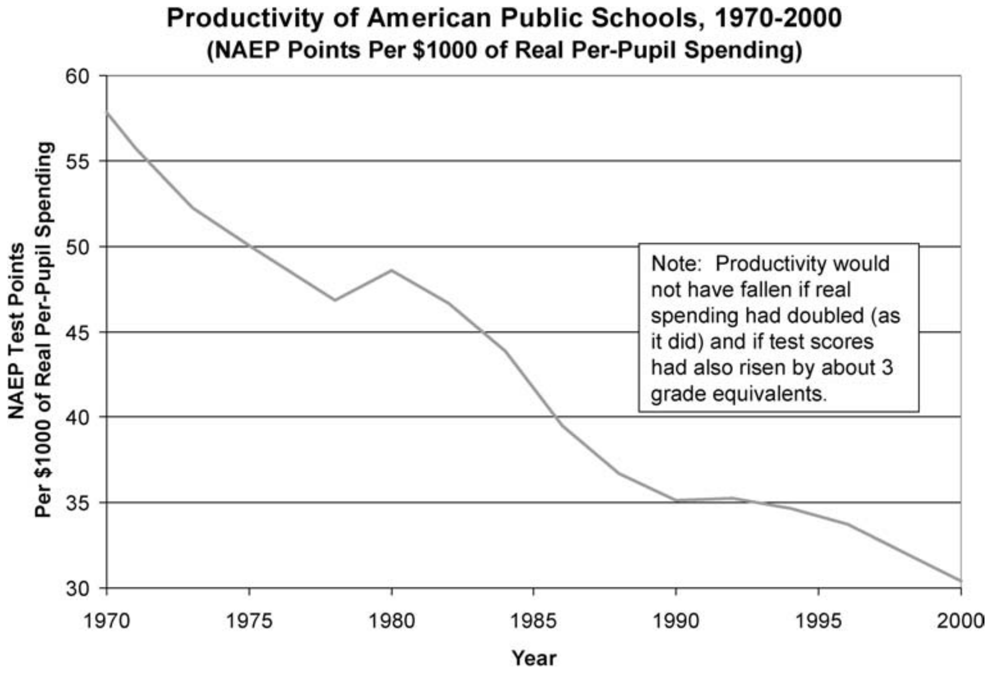

Figure 4. Productivity of American public schools, 1970–2000 (NAEP points per $1000 of real per-pupil spending). Source: Author's calculations using U.S. Department of Education (2003e, 2004) and U.S. Department of Labor (2004) data.

from 1970 to 2000 (U.S. Department of Education 2003c). I have used the Consumer Price Index (CPI) to put dollars of the day into real 2003 dollars (U.S. Department of Labor 2004). In the figure, per-pupil spending rises from $4800 in 1970 to $9230 in 2000.

By dividing the test scores from Figure 2 by per-pupil spending from Figure 3, one gets Figure 4, which gives us estimates of public schools' productivity over time. From 1970 to 2000, productivity fell from 58 to 30 national percentile rank points per $1000. This near-halving of productivity is a decline so substantial that it is hard to ignore.

One might criticize the computations shown in Figure 4 because they take no account of Baumol's "cost disease" argument for why productivity falls in nontraded service industries (Baumol 1967). He argued that the labor in nontraded service industries such as education does not enjoy productivity increases over time, whereas the labor in traded industries is combined more efficiently with other inputs (including human inputs) over time, and the labor in goods-producing industries benefits more from productivity-enhancing technology. Yet, he points out, education must hire people who could work in industries where productivity is rising, so the education industry must pay more and more with each passing year in order to hire workers of the same quality. According to Baumol's argument, the education industry's falling productivity may be beyond our control, an inevitable result of rising productivity elsewhere.

We can take Baumol's argument seriously by assuming that the education sector would have had to pay its workers more (in real terms) with each passing year in order to hire people with the same skills. For instance, we can use an index based on the earnings of professional women in the United States (those with a professional degree such as attorneys, physicians, and masters of business administration) instead of the Consumer Price Index to put per-pupil spending into today's dollars

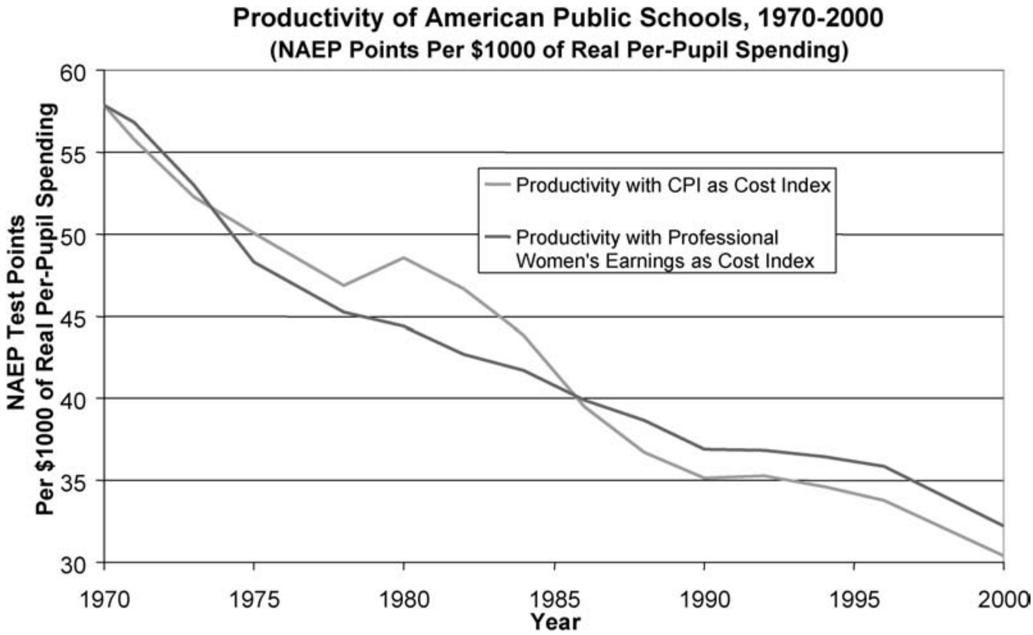

Figure 5. Productivity of American public schools using professional women's earnings to index costs, 1970–2000 (NAEP points per $1000 of real per-pupil spending). Source: Author's calculations using U.S. Department of Education (2003e, 2004) and U.S. Department of Labor (2003) data.

(U.S. Department of Labor 2003). In other words, we will show the productivity of schools over time, holding them harmless for the increasing cost of hiring good teachers. When I use professional women's earnings to generate an index, I am being too generous to schools. That is, I am overadjusting for the rising costs that schools face. The overadjustment occurs because college-educated labor is not the only input into education. Many inputs such as classroom materials and equipment have not had their costs rise significantly over time. I am treating schools as though the costs of all their inputs rose with professional women's earnings. Also, professional women's earnings rose more steeply than the cost of hiring the same quality teacher. This is because I use earnings rather than hourly wages. Professional women worked more and more hours each year, so their earnings grew faster than their wages. Teachers have not worked longer hours with each year. Finally, professional women's earnings rose more steeply than those of constant-quality college-graduate women. This is because professional women are among the most skilled workers in the economy, and the most skilled workers have had earnings gains relative even to other college graduates. In short, I deliberately, substantially overcorrect for Baumol's argument and for other influences on women's earnings opportunities.

Figure 5 shows that, even with this overcorrection for rising input costs, the productivity of American public schools fell significantly from 1970 to 2000. The figure shows productivity falling from 58 to 33 points per $1000 of per-pupil spending. Although one could quibble further about the exact numbers, the estimated fall in productivity is so large that it must be the case that school productivity could be much higher in the United States, as it once was. (You may be surprised to see that, from about 1973 to 1980, the CPI-adjusted line falls more slowly than the female professional earnings line. This is because oil prices figure strongly in the CPI; oil is not an important input in education; and, thus, the CPI-adjusted line understates the decline in educational productivity in the 1970s.)

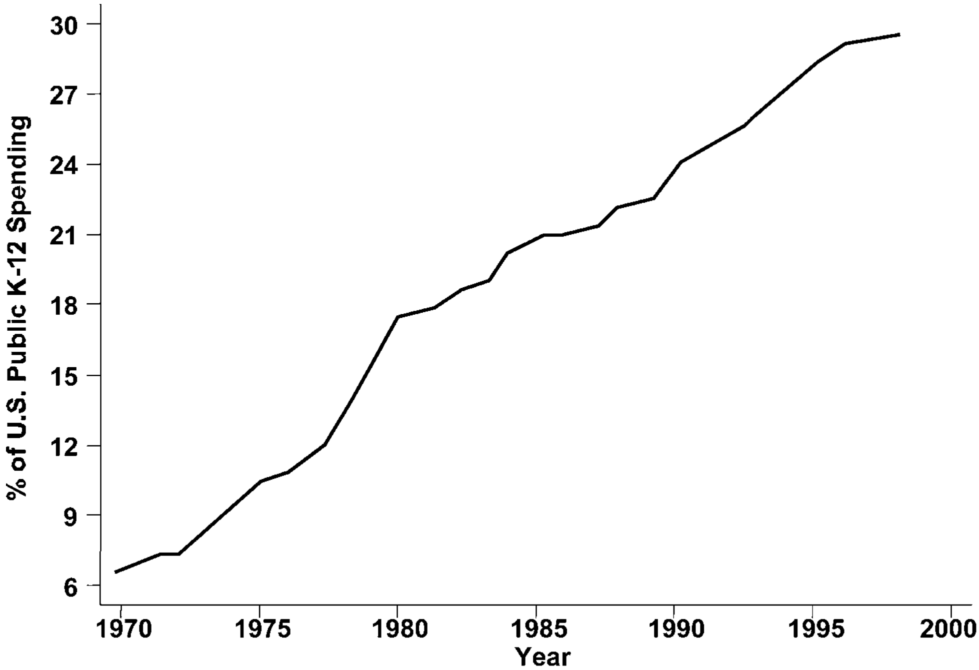

Figure 6. Percentage of U.S. public K–12 spending directed to disadvantaged students (disadvantaged = poor or needing remedial attention). Source: Author's calculations using U.S. Department of Commerce (1975, 1980, 1987, 1991, 1992), U.S. Department of Education (1994, 2000, 2003a, b, d, f), and U.S. Department of Labor (2004) data.

One might ask whether the decline in American school productivity has been caused by our increasing dedication of resources to students who are disadvantaged, either in terms of their family background or in terms of their incoming achievement. I find this question somewhat peculiar because the dedication of resources to such students was based on the theory that they were underserved and that opportunities for achievement growth were therefore high among them. At a minimum, the increasing inputs for disadvantaged children should have raised average achievement by raising their achievement, even if it had no effect elsewhere.

Figure 6 shows that the percentage of American public school spending dedicated to educating disadvantaged children rose from 7% to 30% in the years from 1970 to 2000. Unfortunately, the additional resources did not transform their achievement any more than additional resources transformed the achievement of other students. Figure 7 shows the achievement gap between the average child in the United States and the average disadvantaged child in the United States. Note that the height of the vertical axis is one-half of a standard deviation. The figure shows the achievement gap was just under half a standard deviation in 1970 and remains just under half a standard deviation today.

Now, let us turn to international comparisons of productivity. Figure 8 shows productivity of elementary and secondary schools for OECD countries and a few other countries in a recent year. To construct Figure 8, I used combined mathematics and readings scores from the OECD Program for International Student Assessment (PISA) and mathematics test scores from the Third International Mathematics and Science Study (TIMSS). Both are highly regarded international tests in which students were carefully picked so that countries' samples are comparable. PISA focuses on 15-year-olds; TIMSS focuses on eighth graders, who are about 14 years old. Fourteen- and 15-year-olds are useful because

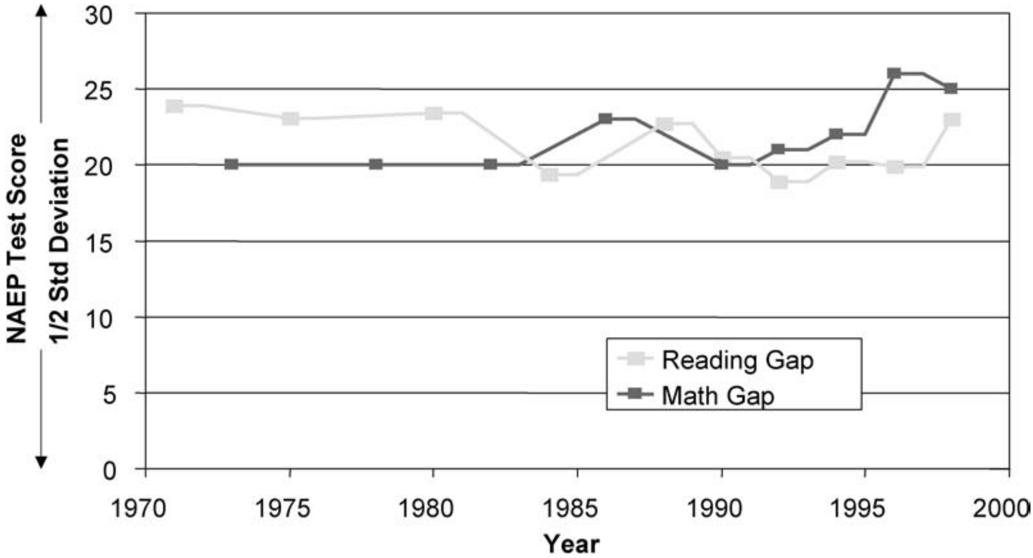

Figure 7. The achievement gap (difference in scores between the average student and students who are poor, black, or Hispanic). Source: Author's calculations using U.S. Department of Education (2003e) data.

nearly all children are still enrolled at these ages in countries that are developed or highly developed. That is, a country will not appear to have higher productivity simply because more of its less able students have already left school. The international public school spending data are for 1999 and are adjusted into U.S. dollars using purchasing power parity (OECD 2002; International Association for the Evaluation of Educational Achievement 1999).

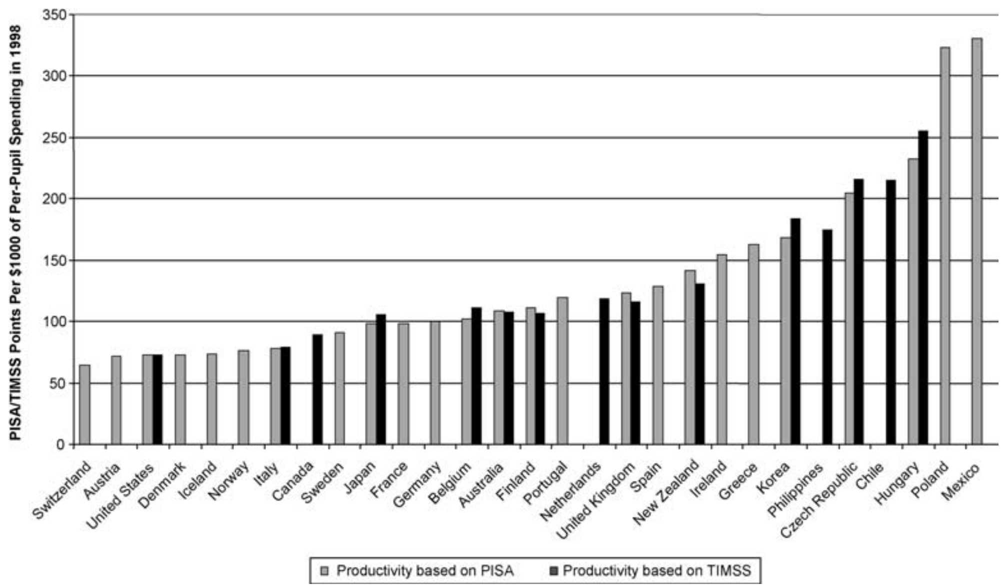

Figure 8. Productivity of public primary and secondary schools (based on PISA 2000 and TIMSS 1999 test scores and 1998 per-pupil spending). Source: Author's calculations using International Association for the Evaluation of Educational Achievement (1999) and Organization for Economic Cooperation and Development (2002, 2003) data.

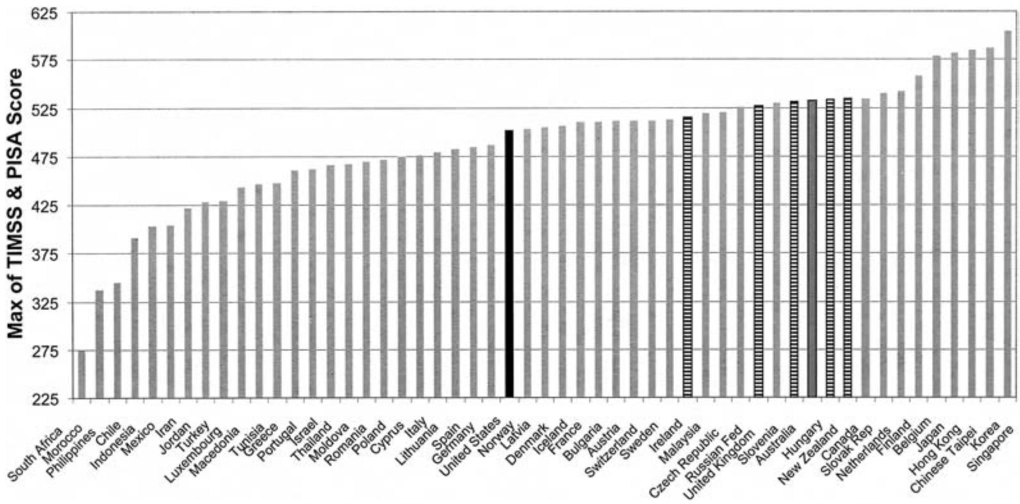

Figure 9. Achievement: United States and rest of world (maximum of PISA 2000 and TIMSS 1999 test scores). Source: Author's calculations using International Association for the Evaluation of Educational Achievement (1999) and Organization for Economic Cooperation and Development (2002) data.

From the international comparisons of public school productivity in Figure 8, it is difficult to conclude that American schools are as productive as they could be. The United States, for instance, is 70% less productive than the United Kingdom, half as productive as New Zealand, and about a third as productive as Korea, the Czech Republic, and Hungary. We really should not be surprised by these differences. Remember that American schools were measured as being at least 79% more productive in 1970. There is no reason to think that other countries are incapable of attaining a level of productivity of which we were once capable ourselves.

One might worry that there are declining achievement returns to every dollar spent and that U.S. achievement is so high that we have low productivity simply because we are on the flat portion of the production function, whereas countries like Hungary are on the steep portion. This is not the case. Figure 9 shows us that U.S. achievement is substantially below that of countries like Hungary. We cannot even complain that international tests are biased against English speakers. U.S. achievement is substantially below that of Ireland, the United Kingdom, Australia, New Zealand, and Canada.

3. The Economics of Why Our Schools Have Mediocre Productivity

If we accept the fact that the productivity of our schools is worrisomely low, then we need economic explanations for why the education sector differs from most other industries in America. After all, most U.S. industries do not have low productivity compared to their international counterparts. Most U.S. industries do not have falling productivity. We know that the higher education industry in the United States does not have low productivity because students from all over the world want to pay to study in the United States. What is it that differs between higher education and K–12 education, that differs between K–12 education and most other industries, that differs between K–12 education today and K–12 education a few decades ago? The prime suspect has to be market power.

The K–12 education sector has not been subjected to many of the reforms and product market pressures that have affected other industries, including semipublic industries such as public housing, utilities, and transportation. With a few notable exceptions, primary and secondary education is highly

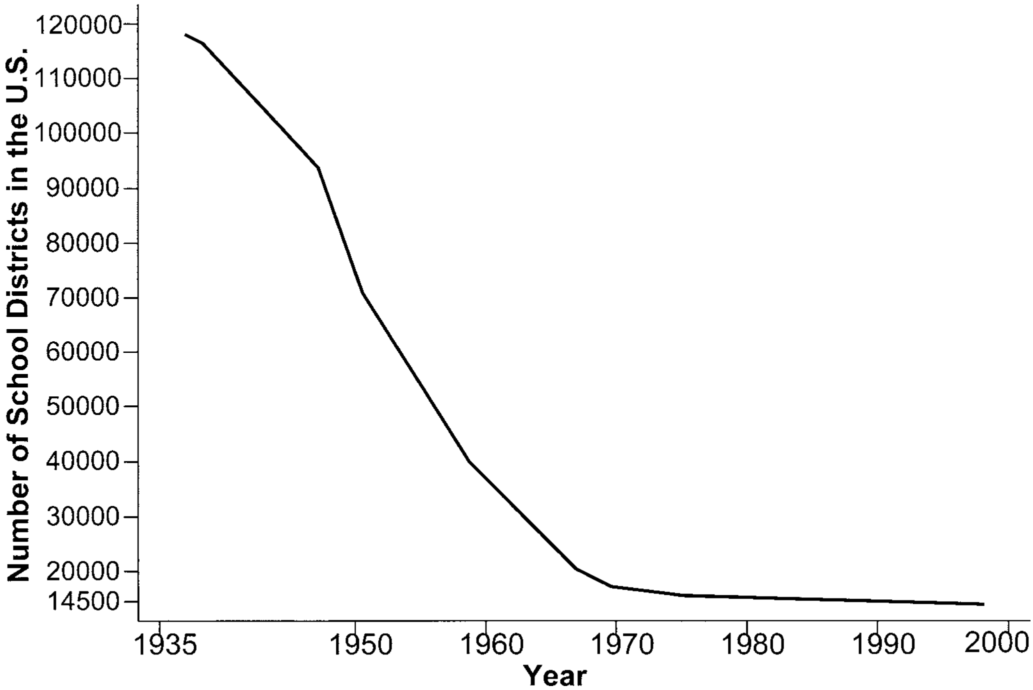

Figure 10. Number of School Districts in the United States. Source: Author's calculations using U.S. Department of Education (2003c) data.

insulated from market pressures, from regulation that attempts to mimic market pressure, and from regulation that attempts to reward performance with pay. Moreover, insulation from market pressure has risen over recent decades.[4] Figure 10 shows that, because of a steady decrease in the number of school districts in the United States, Tiebout competition (the competition that occurs when parents choose a district by choosing a house) has fallen substantially. From 1960 to today, the number of districts in the United States has fallen by more than 85%. Moreover, schools are increasingly insulated from the effects of whatever Tiebout competition there is. This is because the incentive effects of Tiebout work through local property-tax-based school finance, in which a badly managed school drives down local property prices, which drive down its own budget. School finance equalization policies have been enacted with the best of intentions, but they have caused a collateral decline in the effects of Tiebout competition by reducing districts' reliance on local property tax revenues. This is shown in Figure 11. In 1970, the majority of most districts' revenues were local. Today, the majority of revenues come from the state. Figure 11 also shows increasing market power in the key sector upstream of education: the market for elementary and secondary school teachers. This sector had no unionization in 1960. Now, more than 55% of teachers are unionized, and nonunionized teacher contracts increasingly look like unionized ones because of the threat of unionization.

Insulation from market pressures may have direct effects on schools' productivity. There is currently little punishment for schools that spend too much on policies, books, or equipment that are

[4] One might argue that very recent school reforms have increased market pressure and regulation that rewards performance with pay. However, these reforms are so recent and so sparsely applied that they could not yet have affected the national achievement data that appear in Figures 1–2 and 4–9.

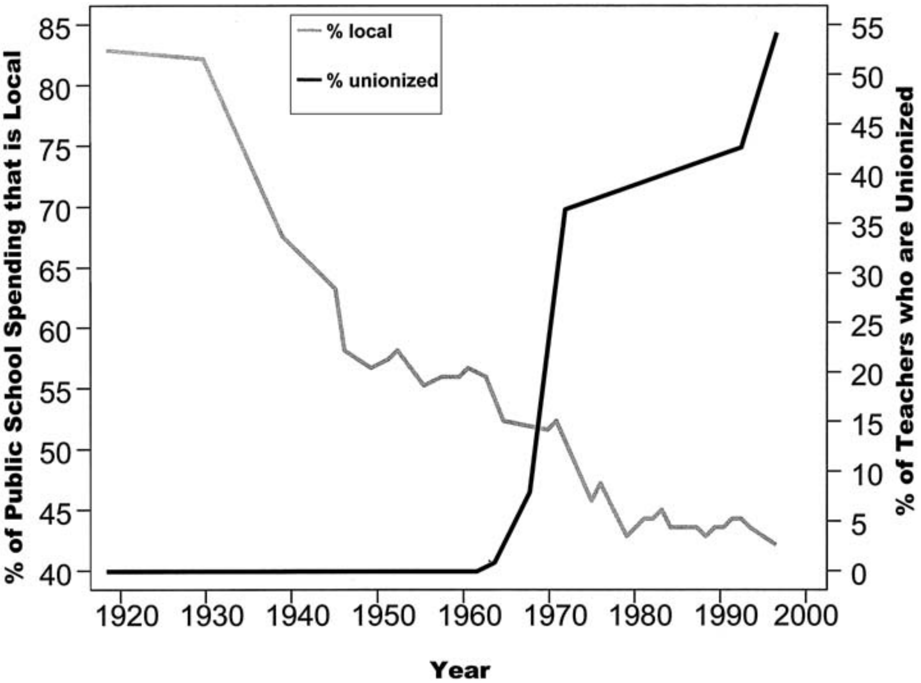

Figure 11. Other indicators of market power (percentage of public school spending that is local and percentage of teachers who are unionized). Source: Hoxby (1996) and author's calculations using U.S. Department of Education (2003c) data.

faddish but have no proven record of raising achievement. There is little punishment for a school with payrolls that are padded with workers who have little to do. There is little reward for a school's managers if they confront their union in an attempt to create more differentiated pay for teachers.

The examples just given are channels by which productivity might be directly affected by insulation from performance-based rewards. But, there are also indirect effects of such insulation. In economies in which nearly every other industry faces growing product market competition or regulation that increasingly attempts to mimic competition, the education sector has become more and more attractive for workers and suppliers who do not want to function in an environment where rewards are based on performance. Roy's seminal model tells us to expect low-performing workers and firms to migrate into industries in which pay is insensitive to performance and, conversely, to expect high-performing workers to migrate out (Roy 1951). Put another way, the education sector needs to have growth in the sensitivity of its rewards to performance merely to stay even with other industries and thereby avoid becoming a magnet for workers and suppliers whose productivity is poor.

4. Policies for Raising School Productivity That are Based on Economics

If we are worried about the productivity of public schools in the United States, there are two basic sources of incentives to which we could turn. The first is accountability—essentially, centralized regulation and yardstick competition based on measured outputs such as test scores. The second is markets. Arguably, European countries and countries such as Japan and Korea use more of both accountability and markets. What we call vouchers in the United States are the norm in most European countries. Indeed, Europeans often find it confusing that what is a "radical" market-based school

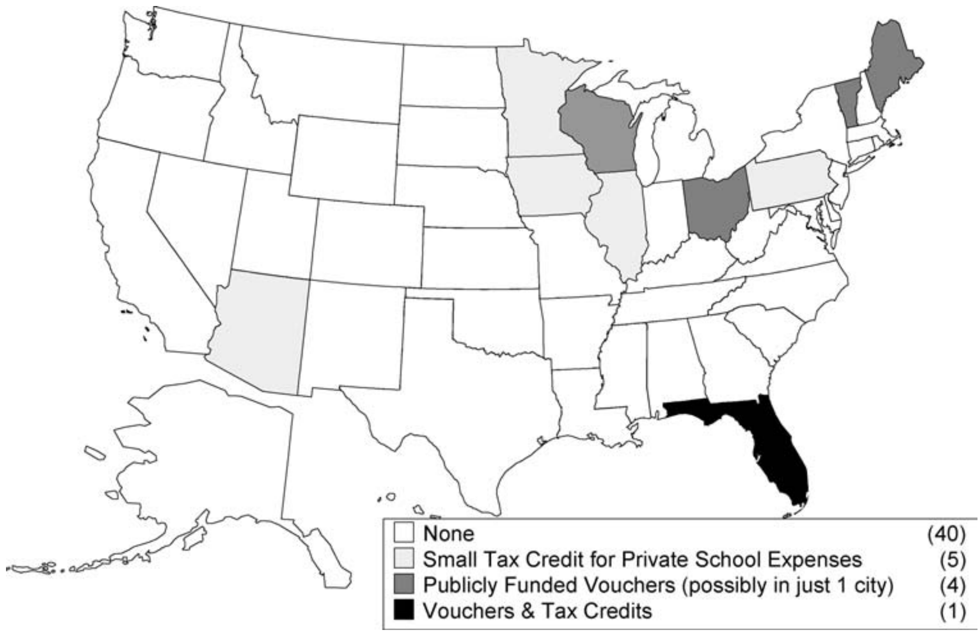

Figure 12. Publicly funded voucher-type programs (2001–2002 school year). Source: Author's calculations using Kafer (2003) data.

choice in the United States is conventional in less market-oriented Europe. In Japan and Korea, private after-school schools play a much more important role than private schools do in the United States. Moreover, centralized examination systems with what Americans would call extremely high stakes are the rule elsewhere in the world. For instance, exit exams are very common.

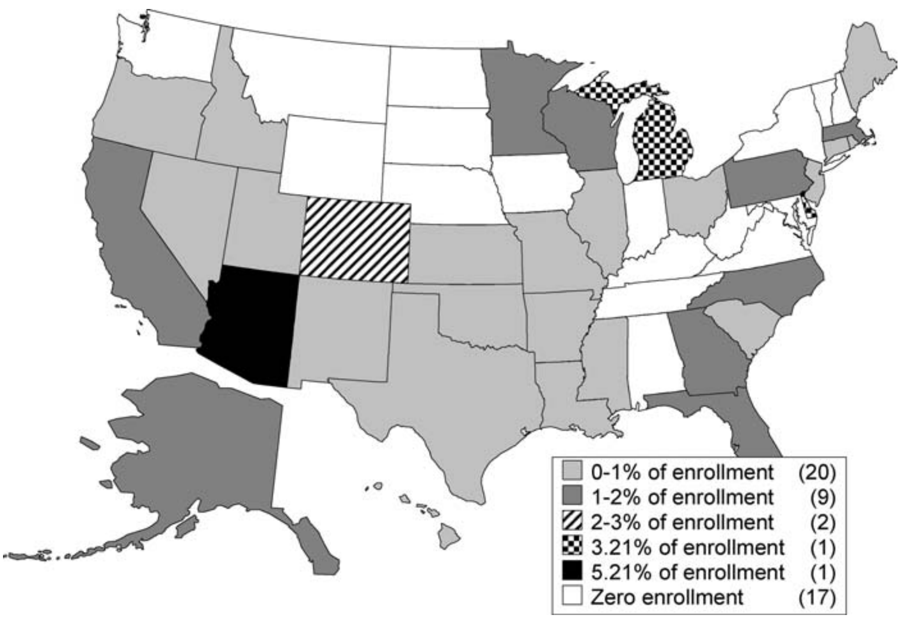

Figure 13. Charter school enrollment in the United States (2001–2002 school year). Source: Author's calculations using U.S. Department of Education (2003b) data.

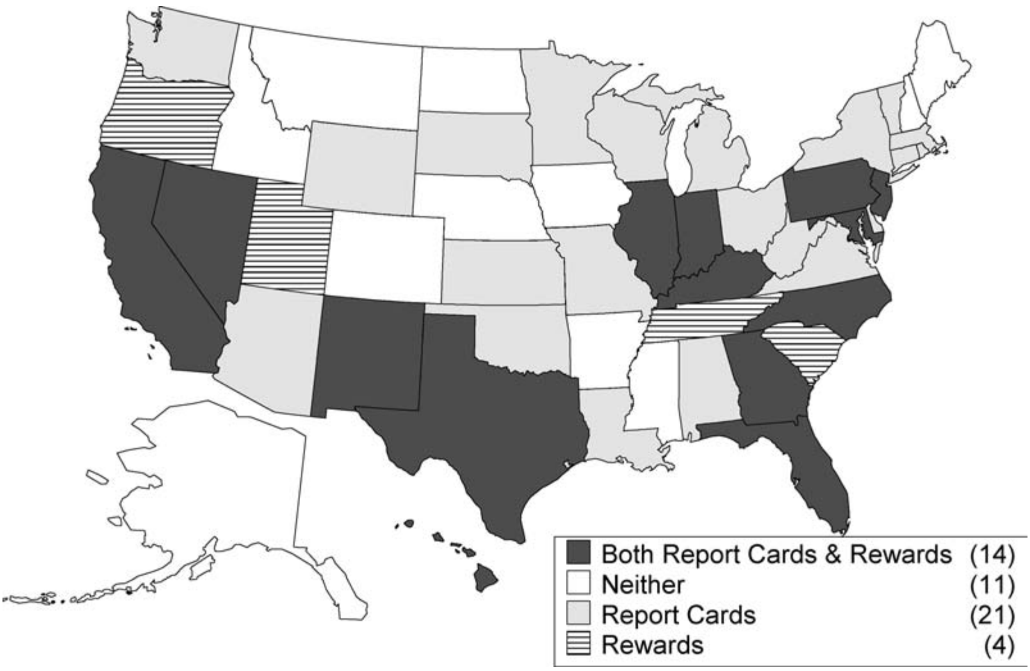

Figure 14. School accountability in the United States (2001–2002 school year). Source: Author's calculations using U.S. Department of Education (2003c) data.

Accountability has its role, especially because it requires the collection of input and output data that can inform many other policies. States have begun to organize schools' administrative data into longitudinal databases that trace students, teachers, and other school inputs. Moreover, some states have begun to link schools' administrative data with data from colleges, employment records, and so on. With such longitudinal databases, we will be able to consider policies that are an order of magnitude more sophisticated.

School Choice and School Productivity

Nevertheless, Americans have always been better at being consumers than at being obedient listeners to someone else's view of their child's achievement and life prospects. Therefore, one would expect the United States to be unusually reliant on markets, as opposed to accountability, to improve incentives in elementary and secondary education. Exactly the opposite is true. There has been so much opposition to market-based reforms in the United States, perhaps because they pose a fundamental threat to established interest groups in education, that we now have substantially more accountability than market reform. The next two maps show how little voucher and charter school activity there is in the United States. Only a few states have anything like vouchers, and most of the states highlighted have made very small efforts: a tax credit of a few hundred dollars or vouchers in a single city's district. Only a dozen states have more than 1% of enrollment in charter schools, and only two states have more than 3% of enrollment in charter schools. To put market-based reform in perspective, consider that home schooling accounts for about five times as many U.S. students as do vouchers and charter schools combined (U.S. Department of Education 2001a, 2004).

In contrast, even before the No Child Left Behind (NCLB) act of 2001, accountability reforms were spreading across the United States. NCLB did little more than codify and reinforce many of

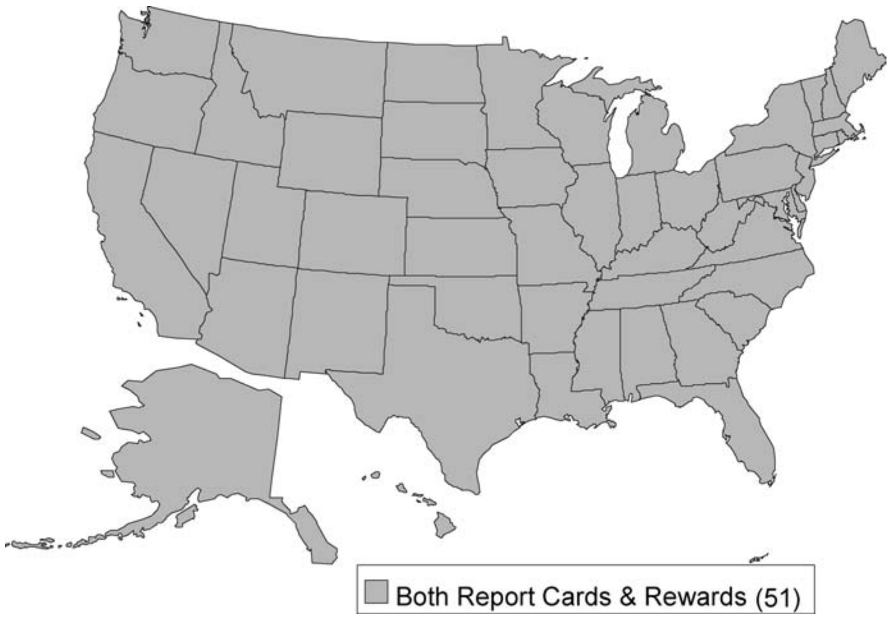

Figure 15. School accountability in the United States (2003–2004 school year). Source: Author's calculations using state compliance with No Child Left Behind legislation.

them. There are three reasons accountability is here to stay. First, state legislatures cannot be expected to supply more than half of the revenues for education without eventually deciding that they need to check on their investments. Just as shareholders ask for audits, state legislatures are asking for accountability. Second, accountability is popular in polls, especially with employers but also with voters in general. Third, accountability is inexpensive. Figure 16 shows the cost of assessment as a share of per-pupil spending in the 25 states with the most comprehensive accountability plans before NCLB. The key feature to note in Figure 16 is that the height of the vertical axis is a mere 1% of per-pupil spending. That is, not even the most expensive accountability plan cost more than about a third of 1% of per-pupil spending (Hoxby 2002a).

Because a lack of market-like competitive pressures is probably the fundamental source of most productivity problems in education, school choice is the reform most prescribed by economics. The crucial feature to remember when discussing school choice is that it has as many manifestations as markets do. No one contemplates primary and secondary education becoming a stereotypical laissez-faire market. Public thought and public money must always be involved if a society is to generate reasonably optimal investments in education. Thus, it is best to think of education as a market that will always be regulated and have its ''prices'' set partly or wholly by the government. The question really is, then, can we give schools better incentives if we use our full knowledge of economics, including our understanding of how to use market competition to design sophisticated regulations? Designing optimal school choice programs is where research is going. There is great interest in programs that can flexibly accommodate peer effects and differences in the cost of educating various students. Here, however, I will not say much about optimal school choice, partly because it is too much in its infancy and partly because it is most useful to establish some evidence about even the crude school choice plans we have.

The key evidence we need to establish is whether public schools raise their productivity when they are faced with conditions that economists would recognize as market-like. By saying ''market-like,'' I refer to choice programs that allow schools to enter, expand, contract, and exit.

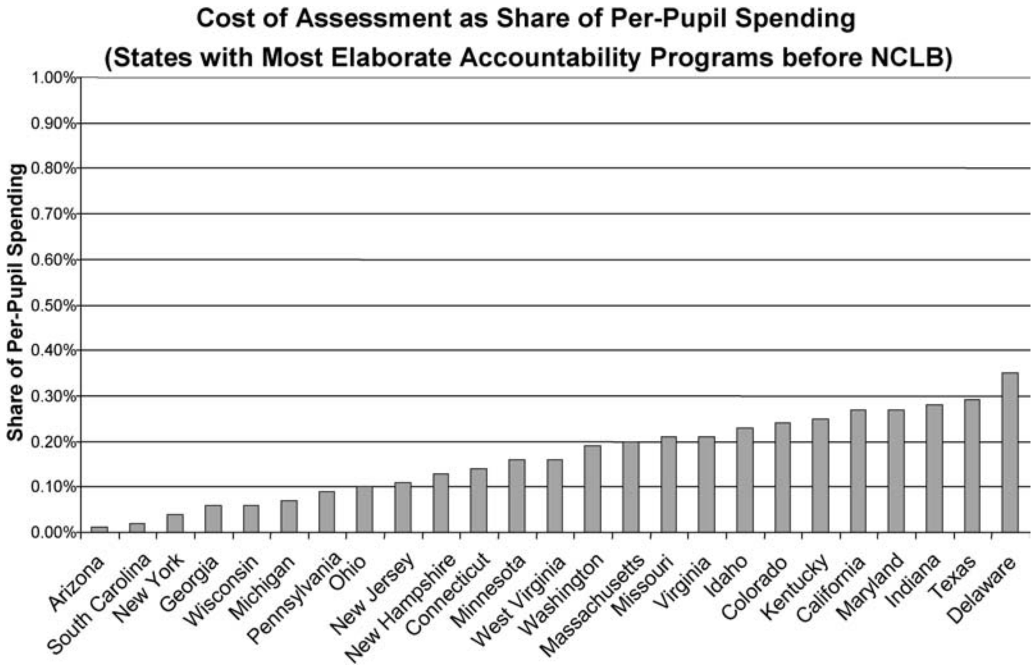

Figure 16. Cost of assessment as share of per-pupil spending (states with most elaborate accountability programs before NCLB). Source: Hoxby (2002b).

A market-like program must also have transfer prices that correspond somewhat to the cost of educating a student: At a minimum, the transfer prices should not be so perverse that they provide schools with the incentive to repel students.

On the question of whether public schools raise their productivity when faced with competition, some of the best evidence comes from Milwaukee, Wisconsin, the only city in the United States with a choice plan that creates significant competition between public and private schools. In Milwaukee, starting in 1998, students within 175% of the poverty line were eligible for a voucher worth 58% of what the public schools in Milwaukee spent. The voucher could be used to pay tuition at a local private school but could not be ‘‘topped up’’ with parents’ funds. For the period in which we will be interested, there was effectively no ceiling on the number of students who could use vouchers, and the vouchers were mainly usable by children in primary grades. Milwaukee contained primary schools in which over 90% of children became suddenly eligible for vouchers. It also contained schools in which only about 30% of children became eligible. The advent of the program in 1998 makes for a nice study because local parents knew about the potential program well in advance, but the program’s actual implementation was a discrete and somewhat unpredictable event because of a court case.

The most credible way to analyze the effects of this policy on the Milwaukee public schools is a straightforward differences-in-differences strategy. Specifically, let us compare the productivity change in Milwaukee public schools that faced the most potential competition with the productivity change in Milwaukee public schools that faced significantly less potential competition and also with the productivity change in urban public schools in Wisconsin that faced no increase in potential competition because they were outside Milwaukee. Keep in mind that I am categorizing schools by the potential competition they faced (that is, the share of students who were eligible for vouchers). In fact, public schools lost only a small fraction of the students who were eligible to leave, probably

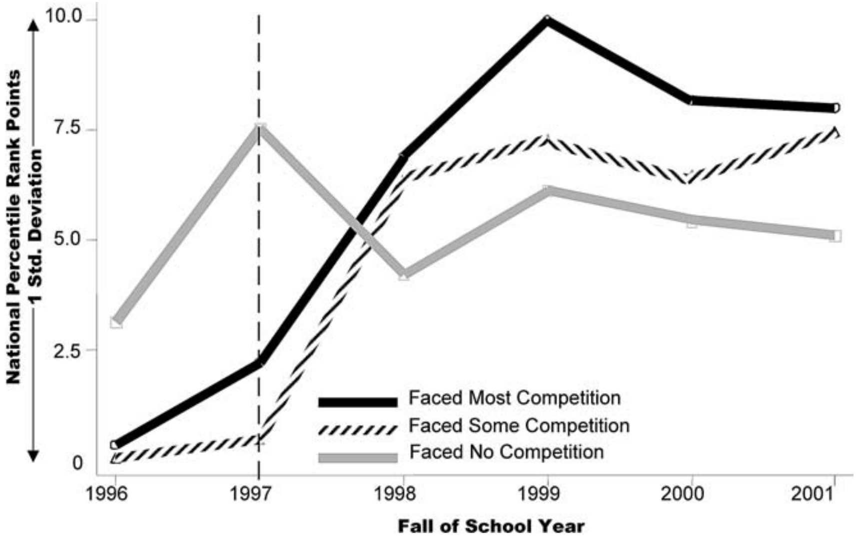

Figure 17. Math scores in Milwaukee and other urban, poor schools in Wisconsin (national percentage rank points, residual from a regression with year and school fixed effects). Source: Hoxby (2003).

because the public schools improved. We can see the differences-in-differences by looking at figures because the changes are quite obvious in the data.

Figures 17 through 19 show, respectively, mathematics, science, and reading achievement in schools exposed to strong potential competition, weak potential competition, and no potential competition. I could show productivity measures instead of achievement, but school spending was similar in the three types of schools, so all of the interest comes from achievement, the numerator of productivity (but see Hoxby 2003 for productivity results). There is a clear increase in the

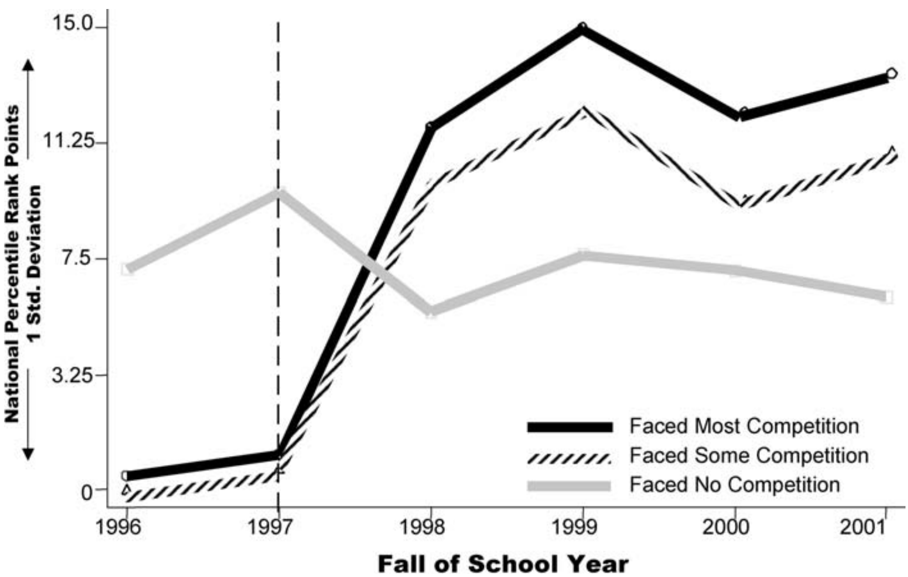

Figure 18. Science scores in Milwaukee schools and other urban, poor schools in Wisconsin (National percentage rank points, residual from a regression with year and school fixed effects). Source: Hoxby (2003).

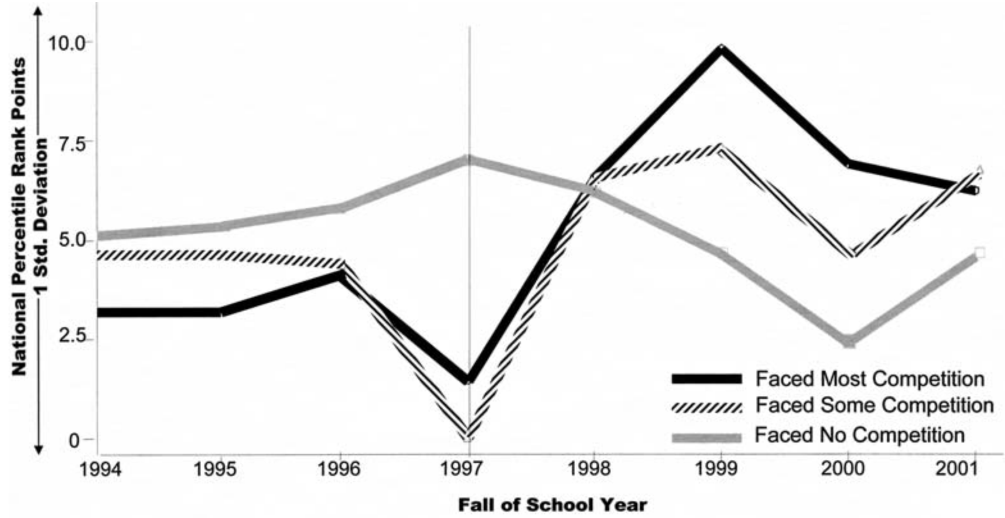

Figure 19. Reading scores in Milwaukee and other urban, poor schools in Wisconsin (national percentage rank points, residual from a regression with year and school fixed effects). Source: Hoxby (2003).

achievement of Milwaukee public school students following the advent of competition, and the improvements are concentrated in the schools that faced the most potential competition. Overall, the differences-in-differences estimators suggest that a school exposed to substantial competition raised its productivity 24% in the three years following the advent of voucher competition (24% is the average across five subject area tests). Milwaukee's gains are some of the largest productivity or achievement gains seen in data on an American school reform.

For those who might be concerned that the increase in achievement could have been caused by the Milwaukee schools' losing relatively low-performing students, it may be useful to have a bound. Even if the voucher-using students had been the lowest-achieving students in the public schools before the advent of the choice program, their departure could not have raised Milwaukee scores by more than a tenth of the amount by which they actually rose. (Remember, because the public schools improved, many students who were eligible to leave did not.)

School Choice and the Sensitivity of Teachers' Pay to Their Performance

Previously, I noted that there are several channels through which schools facing competition might improve productivity. One is that they might increase the sensitivity of their pay to performance. In another study, I surveyed teachers and administrators at most of the charter schools in the United States in an effort to understand how these choice schools manage teachers (Hoxby 2002a). I used survey questions from the government survey of teachers: the Schools and Staffing Survey. By combining my own survey data with the Schools and Staffing data, I was able to construct empirical measures of the sensitivity of pay to a teacher's aptitude, taking on extra duties, and paper credentials. I constructed these measures for the charter schools and for the local public schools from which they drew students.

Charter schools are especially interesting because they are formally public, cannot select students from their applicant pool (if oversubscribed, they must hold lotteries), and must abide by many government regulations. Nevertheless, if charter schools are to survive, they must compete successfully for students. They actively enter markets, grow, shrink, and exit. Interestingly enough, at

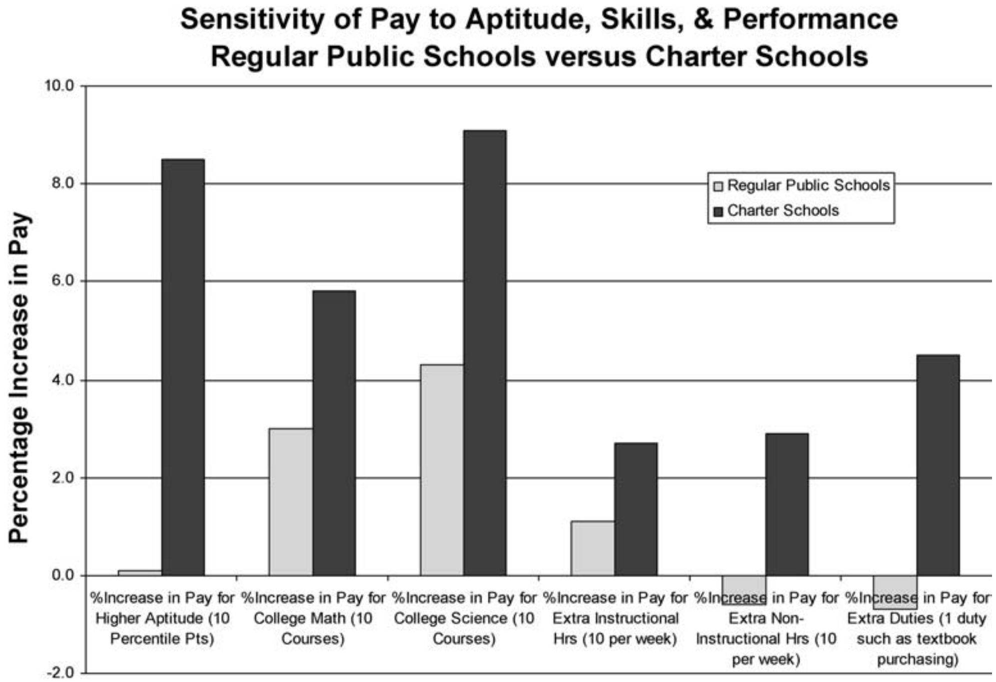

Figure 20. Sensitivity of pay to aptitude, skills, and performance (regular public schools versus charter schools). Source: Hoxby (2002a).

their inceptions, most charter schools simply adopted the pay schedule of the local public school. Thus, if we later observe a difference in the performance sensitivity of their pay, it resulted from their having changed their pay in response to the more competitive environment they face. Charter schools generally did not begin with a different theory of pay.

The results I found bore out the freely offered comments that many charter school administrators wrote on their surveys. They said that they could not keep their best teachers or attract high-performing teachers unless they created differential pay. They said that their experiences had taught them to pay more to teachers who took on extra duties or particularly hard assignments. Parents not only supported the differential pay but had often waged campaigns to get the school to pay its valuable teachers more in order to get them to stay. Largely by a process of trial and error, the charter schools ended up with pay that was significantly more performance sensitive than that of their local public schools. Figure 20 shows these results (in all cases, the difference between the public schools and charter schools is statistically significantly different from zero at the 5% level). Charter schools' pay for higher teacher aptitude (as measured by college entrance examination scores) is substantial: an 8.5% increase for every decile. In contrast, public schools do not pay more to high-aptitude teachers. Compared to regular public schools, charter schools' pay is about twice as sensitive to a teacher's having college preparation in math or science. Public schools do not pay more to teachers who take on extra duties or work extra noninstructional hours; charter schools do.

Recall that the Roy model leads us to expect that, by making pay more sensitive to performance, charter schools would eventually end up experiencing in-migration of high-performing teachers and out-migration of low-performing teachers. Figure 21 suggests that such migration has occurred: Charter school teachers have higher aptitude, took more math and science courses, work longer hours, and take on more extra duties (in all cases, the difference between the public schools and charter

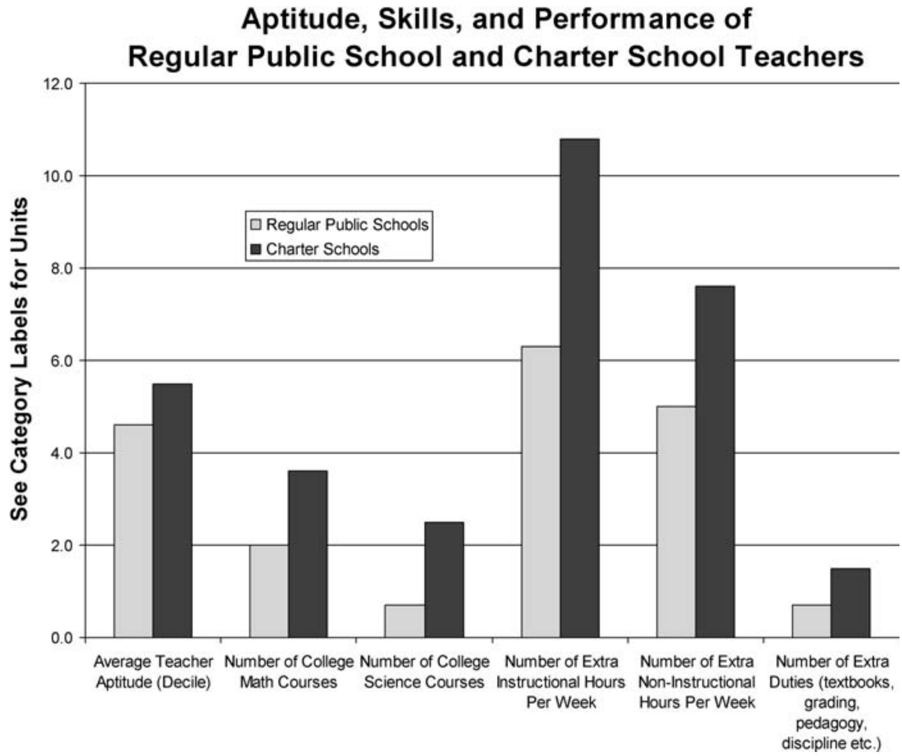

Figure 21. Regular public schools versus charter schools (regular public school and charter school teachers). Source: Hoxby (2002a).

schools is statistically significantly different from zero at the 5% level). Many charter school teachers are drawn not directly from public schools but from private sector jobs. In the U.S. private sector, there are a great many former teachers who are still qualified to teach but who have given up on teaching (the ratio of such former teachers to current teachers is approximately 7 to 1).[5]

5. Teacher Effects and Teacher Pay

Anecdotal experience has long made us suspect that teachers matter, yet the best econometrically identified studies of teacher characteristics show that most characteristics rewarded in teacher contracts (having a master's degree, being certified, experience) have little or no effect on achievement. Recent research has shown, however, that even though such characteristics do not matter, individual teachers do have systematic effects on students. Teacher effects are strong effects. For instance, the difference between the student outcomes generated by the best (top decile) and worst (bottom decile) teachers in an American primary school is seven times the change in student outcomes than the Tennessee Star Study predicts would occur for a 10% reduction in class size (see below for more on the results cited).

Because teachers are potentially so influential, managing them well must be a key factor in any school's reaching its maximum achievable productivity. Teachers need to be hired, promoted, and paid so that people who will be successful become teachers and stay teachers and so that people who

[5] Author's calculations based on Recent College Graduates surveys (U.S. Department of Education 2001b).

will be unsuccessful do not teach. Moreover, addressing teacher pay will address a chronic complaint—that teachers are insufficiently paid and respected. It is unlikely that voters will give substantially more pay and respect to teachers so long as their jobs continue to be some of the least performance-sensitive jobs.

There are two basic reforms that economics suggests will increase the performance sensitivity of pay in teaching. The first is school choice, which I have already discussed. Choice works by creating pressure on schools to reward teaching performance that parents appreciate. We have seen that American charter schools have teacher pay that is more sensitive to performance than do regular public schools. It is also worth noting that the Milwaukee public schools' chief response to school choice was a renegotiation of their teacher contract. Specifically, the contract was rewritten so that teachers could be assigned to classrooms based on their skills rather than strictly on the basis of seniority; teachers could receive extra pay for taking particularly hard assignments; teachers who were chronically absent or delinquent could be counseled out of teaching altogether rather than simply reassigned in the annual "dance of the lemons." In addition, it is interesting to note that every one of the major school management organizations in the United States (the not-for-profit and for-profit firms that contract to run charter schools and troubled public schools) spend a greater share of their budgets on teacher pay than regular public schools do, have differentiated pay that includes large bonuses for teachers who produce unusually good results, and promote teachers to greater responsibility on the basis of performance rather than seniority.

School choice is not, however, the only means by which teacher pay can be made more sensitive to performance. Specifically, the new longitudinal data we have can be used to identify a teacher's effect on almost any student outcome, although contemporaneous student outcomes are the easiest to identify in practice. Using the techniques pioneered for decomposing wages, we can decompose a student's outcome into student, grade, teacher, school, and district effects. That is, we can identify a teacher with the systematic effect that she has on students, given the performance of her students in other classes and grades. Figure 22 shows a straightforward example of the results from one such analysis, taken from Rockoff (2004). The difference in productivity between a top decile and bottom decile teacher in a school is as much as three-quarters of a standard deviation in achievement. The interquartile range in teacher productivity is about a third of a standard deviation in achievement. These are large productivity differences for people receiving identical pay.

Similar analysis of teacher effects has recently been carried out by a variety of other researchers (see, for instance, Horn and Sanders 1998). The studies agree on the following conclusions. Teacher effects are of the magnitude shown in Figure 22; this makes them potentially important. Teacher effects can be estimated with reasonable precision using about four years of longitudinal data. (Getting four years of data is not a problem except with beginning teachers because schools typically have administrative data for several past years, with which they could immediately evaluate their current teachers). A teacher's effect typically improves during her first few years of teaching but thereafter does not grow or decline systemically with age or tenure.

In short, so long as we exercise due agnosticism about data on beginning teachers, results from decomposition of longitudinal data are a potentially rich source of objective information on teacher performance. Although noisy, these performance data are not swayed by personal relationships between a teacher and his or her administrator. Fear of subjective ratings biased by personal relationships has been a major reason why teachers have long resisted performance-based pay. States that can calculate teacher effects could, at a minimum, relay the information to school administrators, who could weigh the information when managing teachers. Other states might be daring enough to devise explicit reward systems based on teacher effects.

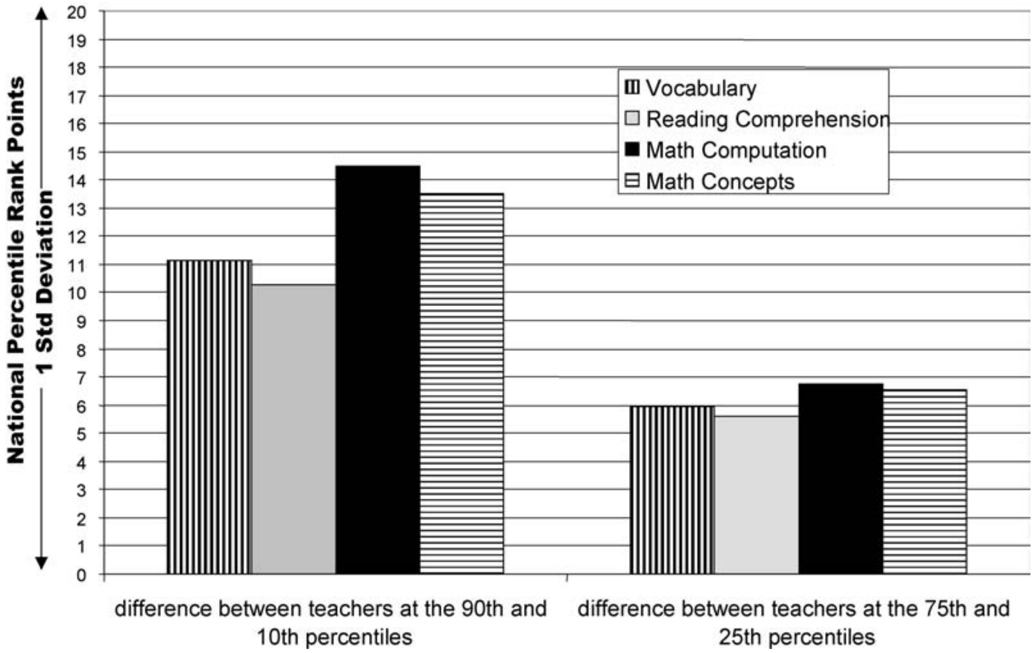

Figure 22. Magnitude of teacher effects in U.S. primary schools. Source: Rockoff (2004).

The best teacher pay systems will probably come from combinations of school choice with estimates of teacher effects. School choice generates information about teachers from parents' observations. These observations are holistic (a plus) but subjective and inexpert (minuses). School choice is best at producing pressure on a group of teachers to perform; a school administrator may find it hard to extract information from choice behavior that allows her to easily address within-school agency problems. Conversely, analysis of longitudinal data generates information about teachers from administrative outcome data. This information is computed in an objective manner (a plus), but it covers only some of the full range of outcomes about which parents and society care. Also, estimated teacher effects are most precise for within-school teacher comparisons. Thus, they are most easily used by an administrator attempting to solve within-school agency problems. In short, school choice and teacher effects estimation are highly complementary policies for increasing the sensitivity of teacher pay to performance.

6. Concluding Thoughts

Many people are wary of changing American schools. They are not thrilled by the performance of their local public schools, yet they are hesitant to change them. Perhaps, they think, we can simply let American public elementary and secondary schools become a sort of productivity "backwater." Perhaps we can think of it as a luxury to have one sector that has low productivity, where professionals do not have to face incentives, that is expensive but just so-so in performance. Unfortunately, although the United States is a rich country, it cannot afford this luxury. This is because education is the upstream industry for the skill-intensive industries that generate our high per-capita incomes. Having the education sector be unproductive will gradually destroy our comparative advantage in the downstream industries we most want to keep. Developing and intermediate countries have plenty of educated individuals. We already see Western firms employing software engineers and consumer service personnel in—for instance—India. With modern technology, firms will increasingly be able to locate the skill-oriented parts of their businesses where the skills originate. In the long run, no country can expect to enjoy skill-intensive industries if its own education sector does not produce skills efficiently.

Practicing the economic analysis of education can be frustrating because research has less sway than politics or media observers' personal experiences. Nevertheless, the tools and insights that economists have developed for use in other areas are incredibly useful in education. Economists need to stay open-minded enough to use all our tools well. For instance, we need to apply not just market logic but also statistical logic (for analyzing teacher effects), the logic of optimal regulation (for designing smart school choice plans), and our knowledge from personnel economics (for designing pay-for-performance plans). If we do, we may be able to devise and evaluate policies that both improve U.S. education and are consistent with social objectives.

References

Baumol, William J. 1967. Macroeconomics of unbalanced growth: The anatomy of urban crisis. *American Economic Review* 57:415–26.

Horn, Sandra P., and William L. Sanders. 1998. Research findings from the Tennessee value-added assessment database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education* 12:247–56.

Hoxby, Caroline M. 1996. How teachers' unions affect education production. *Quarterly Journal of Economics* 111:671–718.

Hoxby, Caroline M. 2002a. Would school choice change the teaching profession? *Journal of Human Resources* 38:846–91.

Hoxby, Caroline M. 2002b. The cost of accountability. In *School accountability*, edited by Williamson M. Evers and Herbert J. Walberg. Stanford, CA: Hoover Institution Press, pp. 47–74.

Hoxby, Caroline M. 2003. School choice and school productivity: Or, is school choice a rising tide that lifts all boats? In *The economics of school choice*, edited by Caroline M. Hoxby. Chicago, IL: University of Chicago Press, pp. 287–342.

International Association for the Evaluation of Educational Achievement, Third International Mathematics and Science Study (TIMSS). 1999. *International mathematics report*. Chestnut Hill, MA: International Study Center, Boston College.

Kafer, Krista. 2003. *School choice 2003: How states are providing greater opportunity in education.* Washington, DC: The Heritage Foundation.

Organization for Economic Cooperation and Development (OECD). 2002. *Program for international student assessment (PISA) 2000. Electronic data.* Paris: Organization for Economic Cooperation and Development.

Organization for Economic Cooperation and Development (OECD). 2003. Tables A1.1 and B1.1. In *Education at a glance.* Paris: Organization for Economic Cooperation and Development.

Rockoff, Jonah. 2004. The impact of teachers on student achievement: New evidence from panel data. Ph.D. thesis, Harvard University, Cambridge, MA.

Roy, A.D. 1951. Some thoughts on the distribution of earnings. *Oxford Economic Papers* 3:235–46.

United States Department of Commerce, Bureau of the Census. 1975, 1980, 1987, and 1991. *Census of governments: 1972, 1977, 1982 and 1987 data.* ICPSR edition. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor].

United States Department of Commerce, Bureau of the Census. 1992. *Census of population and housing, 1980: summary tape file 3F.* Washington, DC: United States Department of Commerce, Bureau of the Census [producer], 1991. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

United States Department of Education, National Center for Education Statistics. 1994. *School district data book: 1990 census school district special tabulation.* Arlington, VA: The MESA Group.

United States Department of Education, National Center for Education Statistics. 2000. *Elementary and secondary general information system: public school district universe data, 1969–1986* [Electronic files]. ICPSR Version. Washington, DC: United States Department of Education, National Center for Education Statistics [Producer]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [Distributor].

United States Department of Education, National Center for Education Statistics. 2001a. *Home schooling in the United States 1999.* Washington, DC: National Center for Education Statistics.

United States Department of Education, National Center for Education Statistics. 2001b. *Recent College Graduates. 1974–75, 1976–77, 1979–80, 1983–84, 1985–86, 1989–90 data.* Electronic data. Washington, DC: National Center for Education Statistics.

United States Department of Education, National Center for Education Statistics. 2002. *National education longitudinal study 1988/2000*. Washington, DC: National Center for Education Statistics.

United States Department of Education, National Center for Education Statistics. 2003a. *1970 census fourth count (population): school district data tape*. Electronic file. ICPSR version. Washington, DC: United States Department of Education, National Center for Education Statistics [producer], 1970. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

United States Department of Education, National Center for Education Statistics. 2003b. *Common core of data: public elementary/secondary school universe survey, 1986–87 to 2000–01 school years*. Electronic files. Washington, DC: United States Department of Education, National Center for Education Statistics, 2003.

United States Department of Education, National Center for Education Statistics. 2003c. *Digest of education statistics, 2001*. Washington, DC: National Center for Education Statistics.

United States Department of Education, National Center for Education Statistics. 2003d. *Elementary–secondary school system finance data files for fiscal years 1992 to 2002*. Washington, DC: United States Department of Education, National Center for Education Statistics.

United States Department of Education, National Center for Education Statistics. 2003e. *National assessment of educational progress*. Washington, DC: National Center for Education Statistics.

United States Office of Education. National Center for Education Statistics. 2003f. *National institute of education special tabulations and 1970 census fifth count data file*. Computer file. ICPSR version. Washington, DC: AUI Policy Research [producer], 1981. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2003.

United States Department of Education, National Center for Education Statistics. 2004. *Digest of education statistics, 2002*. Washington, DC: National Center for Education Statistics.

United States Department of Labor, Bureau of Labor Statistics. 2003. *Current population survey 1964–2003 data*. Washington, DC: United States Department of Labor.

United States Department of Labor, Bureau of Labor Statistics. 2004. *Consumer Price Index*. Washington, DC: United States Department of Labor.