Paying Teachers for Performance: Incentives and Selection

Edward P. Lazear

Hoover Institution
and
Graduate School of Business

Stanford University

Preliminary

August 19, 2001
Abstract

There has been much discussion of tying teacher compensation to performance. The usual rationale is that this will provide teachers with appropriate incentives and that students will benefit as a result. That argument is explored in this essay. It is argued that teacher pay is likely to affect outcomes as much through the selection mechanism as through any incentives generated by performance-based pay. Furthermore, defining performance is key and has been a topic that has not received sufficient attention. The choice of the metric affects dramatically the distribution of benefits among students. Finally, teachers’ desires for smaller classes are examined and related to the existence or lack thereof of competitive market pay.
“Pay teachers on performance,” is a rallying cry of education critics and the profession’s most vocal detractors. The usual notion is that teachers, in part as a result of the tenure system, have become complacent and can be motivated by incentive pay. Incentives surely matter. There is no need to assume that teachers are lazy. Pay may motivate teachers to move in one direction over another or to forego some marginal leisure activities if the reward is sufficiently high. Although incentives are discussed below, it is easy to overemphasize the importance of motivation and to neglect other rationales for performance based pay.

Also mentioned sometimes is that pay for performance will result in low pay for the poorer teachers and high pay for the better ones, which might encourage the right pattern of retention and turnover. This point, although less frequently emphasized, deserves significant attention. However, arguments about selective turnover are tenuous at best. Because poorer teachers also have poorer outside opportunities, pay dispersion would have to be increased considerably to induce the appropriate turnover. More important, perhaps, is using pay to create a larger pool from which to hire. If the selection of applicants is greater and if teachers can be evaluated for longer periods of time before tenure is granted, then pay may be a tool for bringing about improvement in educational output. A primary theme of this essay is that selection as well as motivation is a rationale for performance based pay. Additionally, the definition pay for performance requires much more specificity before it can be evaluated in any appropriate way. Other pay and incentives issues are also discussed below.
Sorting and Incentives:

It is instructive to consider an example from a completely different industry, namely the auto glass business, where individuals install windscreens in automobiles. In Lazear (2000), an analysis is presented of a company (Safelite) that switched from paying hourly wages to paying on the basis of performance. The results were striking. Virtually overnight, productivity increased by 44%. Half of the increase was attributable to incentives: The average individual worked harder after being paid on the basis of output and this accounted for about a 22% rise in productivity. More surprising was that the work force in place about one year after the initiation of the plan was quite different than the one in place when hourly wages were paid. Specifically, the new workers were about 24% more productive than the workers they replaced. The two effects combined to produce an increase in productivity of 44%.

Of course, teaching is quite different from installing windscreens. First, and perhaps most important, the output in education is more difficult to define and to measure. Second, output is multidimensional so the metric in education would have to be some weighting of various components. Third, what is best for one student is not best for another. All of these factors cause the industries to differ. That notwithstanding, the general point is valid. Incentives matter, but so does selection. The new pay scheme at Safelite induced a completely different applicant pool than previously encountered and the company benefitted as a result.

Incentives

If performance pay matters for incentives, it is first necessary to understand that the measure of
performance is crucial because it so determines what is produced.¹

Measuring Output:

There are a number of ways to measure the performance of a school or of its teachers. The most frequently-used measure is student test scores, but there are many others that have been referred to in the past. Literacy rates among students of a certain grade level, high school completion rates, and the proportion going on to college are all indicators of student achievement, but they indicate different things. The measure of output chosen can affect whether a school is viewed as good or bad, and can also affect the strategies adopted by a particular school.

Suppose that a school is judged on literacy among its eighth grade school population. That school will attempt to get every child to some basic level of reading, but it will not have incentives to focus on helping particular students excel in reading. Nor does the school have strong incentives to worry about mathematical skills, because those skills are not measured by the standard literacy tests.

Suppose instead that we judge a teacher or school by the proportion of students who graduate from high school. This will cause the school to focus on those students who are right on the margin between staying and dropping out. The poorest students will be ignored because they are unlikely to be swayed into remaining in school by policies that a school can implement. The best students are also likely to be ignored because they are not at a high risk of dropping out. As a result, schools target a particular part of the student population, in this case at the low, but not the lowest part of the achievement distribution.

¹Much of what follows is excerpted from Lazear (2002).
At the other extreme, one can imagine basing an assessment of school performance on the proportion of students who go on to college, or even on those who go on to elite colleges. Compensating on the basis of the proportion going to Oxford or Cambridge would influence schools to focus on the best students in the class, perhaps to the disadvantage of those in the middle or bottom of the class.

Another measure of school performance looks at the bottom line: How do schools affect subsequent earnings or the occupational distribution of their students? For many purposes, this is exactly the right measure of school performance. The problem with this measure is that it is very difficult to obtain, and even when possible, it comes only after a very long lag. By the time the students have earnings to report, the school may have already changed its policies many times over.

Other Issues of Accountability:

Few policy makers, including those who are part of the education community, would actually argue against accountability. The main problem, however, is defining appropriate standards for accountability and knowing which factors to take into account.

Suppose that we ignore the arguments of the previous section and simply institute a uniform test against which schools are compared. A number of problems remain.

First, schools that have students from wealthier and more educated homes are likely to obtain higher test scores even if the school in question is contributing no more to a child’s education than other schools in the sample. In other words, demographic characteristics of the student population are likely to have significant effects on test scores. Presumably, society is interested in added value associated with education and not merely a certification that schools have been able to attract a bright group of entrants.
Local to my home institution of Stanford University, the Palo Alto schools are known for the high test scores of their students. Is this because the schools are doing a good job, or is it because a large proportion of the students are the children of Stanford faculty or other professionals?

Second, some schools have more resources to work with than others. It is unreasonable to expect that schools operating on a per-student budget that is half that of other schools will be able to produce the same quality of educational output. Going back to the previous argument, if nothing else, it will be much more difficult to attract quality teachers on a smaller budget than one that allows higher teacher wages to be paid. The richest districts will have the ability to cherry-pick among teacher applicants.

As a potential solution, it is possible to look at gains in scores. That is, a district that has the good fortune to have high-achieving students and/or deep pockets can be asked to make the same percentage gains in achievement as those of poorer or less well-positioned schools. Thus, a school with an initial average test score of 600 would be required to move its students to 660 when a school that had initial test scores of 500 was required to move its students to 550. Although something along these lines sounds fair, it penalizes schools that have done well in the past. For example, consider a school that has done the best possible job for its students. As a result, test scores are high, students are happy, graduation rates are among the highest in the nation. It may be very difficult to add to its stellar record, simply because all the appropriate steps have already been taken. Such schools would be penalized by a system that looked for change rather than levels of performance.

What, then, is the solution? Conceptually, it is necessary to hold constant demographic characteristics of the student and resources available to the schools. This requires comparisons between similarly-situated schools. Although it is straightforward to do this statistically, there remains one major
problem. The differences that one observes in test scores across schools are as likely to reflect unobserved differences in school or student characteristics as they are to reflect differences in the actual performance of the school itself. For example, in comparing a school in one part of a wealthy town to another part of that wealthy town, the differences in the test scores obtained by the students are likely to be small. Those small differences may reflect differential performance of the two schools, or small differences in the characteristics of the students who live in different parts of the town. Thus, accountability, although important, is not easily implemented. It can be an important part of the policy equation, but must be implemented by those who have a sophisticated knowledge of statistical methods as well as educational practices.

One way around the problem is to allow more choice for students. This does not necessarily require private schools, but allowing mobility between schools will provide good signals of how well a school is actually performing. If students tend to move from School A to School B, then one can conclude that students feels that they get more out of School B than School A. Although it is still possible that other factors are involved, this is probably the best indicator that School B is doing a better job than School A in educating it students. Thus, students vote with their feet and provide incentives to schools and their teachers to perform or mimic those who succeed in attracting students.

Incentives to Governments: Central versus Local Administration of Schools

In the United States, most of the money going to education comes from the state and not from the federal government. Indeed, the largest federal program accounts for only about 14 billion dollars nationwide. The state education budget in California alone is triple that figure. Thus, it is difficult to expect
the federal government to do much to influence education policy, which is made primarily at the state and local level. There have, however, been some attempts to use federal muscle to influence local educational decisions. Before analyzing such policies, it is important to ask whether federal policies are appropriate, or if they are better left to the state.

The United Kingdom is smaller than the United States, but also diverse in its population and region. Thus, at least some of the issues that confront the U.S. also confront the U.K.

There are at least two arguments against allowing the central government to play an important role in educational policy. First, when the central government is guiding the nation’s education policy, the stakes are much higher than is the case when localities are guiding the education policy. The impact of educational policy at the national level is far greater than that at the local level. For example, if all schools were required to use the same textbooks for a particular subject in a particular grade, the profits associated with inducing a central level administrator to choose a particular publisher’s book would be far greater than those associated with inducing any single local administrator to do the same. As a result, lobbying pressure would be far greater if power were centralized and central agencies more subject to capture than localities.

Second, even if central authorities make what everyone believes to be the right decision, assigning decisions to the central level creates an extremely risky situation. When decisions are made at the local level, a wrong decision does not result in a nationwide disaster. When decisions are made at the central level, making a mistake in education policy does not just affect the students in a given locality, but could be devastating to an entire generation of the nation’s students. Furthermore, there would be no way for individuals to escape such a policy. If a locality made a bad decision, parents who were seeking to improve the welfare of their children at least would have the option of moving out of that locality to another public...
school setting. Central requirements would allow no such movement, and competition would be stifled between districts.

In the United States, a third argument holds. Unless the central budget for education is increased dramatically, it is unlikely that policies at the central level will have the teeth required to get policies implemented. States and localities will be willing to comply with central policies only insofar as the amount that will be lost by failing to comply exceeds the costs imposed on the districts by the policies themselves. Thus, localities will be willing to make small changes, but the central government’s ability to influence local policy will be greatly limited by small amounts being transferred to localities.

Teacher Incentives and Class Size:

Teachers are almost universally in favor of smaller class sizes despite the fact that the evidence on the effectiveness of class size reduction on educational output is far from unequivocal. There are two possible explanations for teacher sentiment. One possibility is that the evidence on class size is simply incorrect or perhaps interpreted inappropriately. Another more cynical possibility is that teachers’ reasons for wanting class size reduction is that smaller classes make for easier jobs and, other things equal, teachers prefer lighter to heavier workloads. Each argument is considered in turn.

A fallacious interpretation of the evidence

The usual way to deal with problems in the public sector is to assume that increasing expenditures will solve the problem. In education in the United States at least, the focus has been primarily on reduction
of class size. But there is a large literature that suggests that altering class size has no effect on outcomes.

There are a few studies, however, that do find important class size effects. Krueger (1998 and 1999) and Angrist and Lavy (1999) find that reducing class size has beneficial outcomes. Since reducing class implies very large costs to taxpayers, it is important to understand the data before a blanket policy of class size reduction is introduced.

Why do some studies find effects whereas others do not? The answer to the puzzle is that classroom education is what economists call a “public good.” That is, one child can benefit from a teacher’s instruction at the same time that another benefits from the same instruction. If both listen attentively to the lesson, then both can obtain the human capital being provided by the teacher.

The problem in the real classroom is that students do not always listen attentively to the teacher’s instructions, and in a public good setting of this sort, when one child acts out, he reduces or eliminates the instructional component of that moment in the classroom for all the other children in the class. Of course, educators are well aware of this, and it is the reason why pre-school children are placed in smaller classes than are most college freshmen. (In fact, as an undergraduate at UCLA, I was in a class with two thousand other students who watched the professor on television in four rooms that sat five hundred each.) Much of the controversy can be eliminated once it is understood the education is a public good subject to negative spillovers from each of the students. The difficulty arises from the allocating of better students to larger

---

1 See for example, Hanushek 1998, who finds little evidence that anything, including class size reductions, matters. Also Coleman and Hoffer (1987), and Coleman, Kilgore and Hoffer (1981) report that Catholic schools with larger class sizes produce better students than public school classes against which they are compared.

class sizes. As a result, even if class size effects were important, they would be difficult to observe in data from the real world, because the better-behaved and presumably more able students are in large classes. The less well-behaved, and presumably younger students, are in smaller classes. When researchers examine large classes, they find educational outcomes are sometimes better in the large classes than in small classes. The reason, however, is that students are not randomly assigned to these classes, but are sorted according to their ability with the better students being in the large classes. So, the failure to observe class size effects may simply be a result of ignoring the fact that larger classes are associated with better students.

To see this, consider an extreme case. Advanced placement students are often found in very large classes because these students are relatively well-behaved and sit quietly through instruction. At the other end of the spectrum are students with behavior disorders who tend to be placed in small classes. A naive analysis would find the large classes with the able students out-teaching the smaller classes with the less able students. This does not mean that reducing class size for the able students would not have beneficial effects, nor does it imply that increasing class size for the disruptive students would not have harmful effects. But it does mean that such effects will not be observed in studies that cut across class sizes.

Does this mean that reducing class size is the solution? Not at all. In fact, the best studies tend to find that to the extent class size effects are important, they are not universal. Disadvantaged children, either as a result of economic status or learning ability, are most likely to benefit from smaller class sizes. Similarly, younger children are also likely to benefit from smaller class size.

Finally, a point made by Eric Hanushek is that teacher quality is an extremely important determinant of educational performance. In fact, in his Texas study, Hanushek finds that putting a child in a good teacher’s classroom is much more important in terms of affecting student learning than almost any other
Because the majority of private schools are Catholic, the reality is actually the opposite of what would be predicted. Teachers in public school receive higher salaries than those in private schools, despite the larger class sizes associated with private schools.

This is shown in Lazear (2000a).

Teachers want reduced workloads

As many of us who are teachers know, teachers may have preferences about classroom composition. A primary consideration of teachers is class size. A cynical view is that smaller class size is a teacher benefit, not a student benefit. It is useful, therefore, to ask whether teacher interests are aligned with the interests of their students.

In an efficient labor market, teacher salary would not be independent of class size. This is the point of Antos and Rosen (1975) who estimate teacher hedonic wage functions. In a competitive labor market, teacher salaries adjust to incorporate the marginal teacher’s willingness to pay for smaller class size. Teachers who teach in larger classes receive higher salaries to compensate. As long as the market adjusts salary efficiently, there is no tension between student preferences and teacher preferences. If teachers prefer smaller classes than students, then they will give up enough earnings to compensate for the smaller class size. This will always result in efficiency. Increasing the number of teachers implies an increase in costs because teacher salary is positive. Increasing the number of teachers also implies a decline in cost per teacher because class size is reduced and teachers (at least on the margin) are willing to give up wages for smaller classes.

In reality, teacher salaries are not subject to the same competitive forces as most private sector

---

4 Because the majority of private schools are Catholic, the reality is actually the opposite of what would be predicted. Teachers in public school receive higher salaries than those in private schools, despite the larger class sizes associated with private schools.

5 This is shown in Lazear (2000a).
salaries because most teachers are hired by the state. It is clear that teachers must be paid at least as much
as they can receive in their alternative occupation, but there is no constraint on the top. Unlike private
firms, states do not go out of business when their costs are too high. As a result, teachers' salaries need not
adjust to non-monetary job attributes as they would in a competitive market. Consequently, teachers' interests
are not necessarily efficient; they may seek smaller classes even when class size reduction is not
cost effective.

Even if markets fully internalize the non-monetary tradeoffs in class size, there is another reason why
teachers as a group might prefer smaller class sizes. A reduction in class size, coupled with compulsory
schooling, implies an increase in the demand for teachers. The recently mandated class size reduction in
California meant a significant increase in demand for teachers. Some was probably reflected in increased
salaries, some in ease of obtaining a teaching job and some in a lowering of standards for teachers. It is
not surprising that an organized group would push for policies that increase the demand for its services.\(^6\)

Incentives in Private and Public Markets:

Private schools charge a price that depends on the quality of the education that it produces. While
other factors may affect the price (such as charity to needy students), a school that competes in a private
market is constrained to charge a price that is commensurate with the quality of the education it produces.
Even though schools are non-profit organizations, the revenue that is generated by an improvement in
quality of the product goes directly to the school in question. This, in turn, makes the jobs of administrators

\(^6\)If school districts' budgets were fixed, then the argument is less clear. Incumbents might lose
as schools were forced to lower wages and substitute quantity for quality in teacher hiring.
easier. Also, workers in non-profit organizations may capture at least some of the returns to higher net revenue. The wages of administrators and teachers are likely to be related to the revenue of the school.⁷

Public schools are different. If a district offers higher quality schooling, the value of land in that district rises, but the school is unlikely to capture much of the revenue from improved quality. Even if tax revenue goes up as a result of increasing land prices, only a fraction of the tax revenue may be passed to the schools in general and the school that actually generates the additional revenue is unlikely to capture much of the return. The more centralized is the school system, the less direct impact on individual schools.⁸ As a result of the centralized revenue pool and decentralized work decisions, there may be more incentive slippage in the public schools than in the private ones.⁹

Wages may be somewhat responsive to firm profitability and to individual performance in private schools, but are almost totally divorced from productivity considerations in the public schools. Wage setting is done centrally and the group of government workers, to which teachers and school administrators belong, have wages that are based on rigid grade and step scales. Finally, private schools can go out of business more easily than can public schools, although this should not be overemphasized because public

---

⁷In a completely competitive labor market with free mobility, there is no necessary relation between profits and wages. To the extent that there is an attachment that is firm specific, there can be some ex post rent sharing.

⁸States differ in this regard. In California, property taxes are collected and distributed at the state level subject to rigid egalitarian formulas. In Illinois, taxes are county based and revenues that flow to schools are more directly tied to the specific county’s tax revenue.

⁹See Hoxby (1998). Also, it is conceivable that the same problem could plague private schools that were part of a chain. Just as a single McDonald’s fails to take into account the effect of its actions on McDonald’s brand name, so too might one branch of a large system of private schools ignore the impact of its actions on the value of the system.
schools can and frequently do close their doors.

The consequence of these differences is to create incentives in public schools that differ from those in the private ones. It can be argued that limitations on the ability to take compensation directly in the public sector induce officials to take their compensation in other forms. Just as contractors shade quality when they are paid a fixed price for a construction job, so too might school administrators shade effort when they are paid a fixed wage for a job that is independent of quality. If the wage in the public sector is less responsive than the wage in the private sector, then incentives to perform are likely to be weaker in public schools.

It is possible to overstate the differences in incentives between public and private schools, which share some common attributes. First, there is a market for successful administrators and teachers that operates in both public and private schools. Second, the incentives to keep a constituency happy are present in both sectors. Parents and students can make their displeasure known as readily in the public sector as in the private one. Parental unhappiness may have less impact on school revenue in the public sector than in the private sector, but this does not negate the fact that teachers and administrators must deal with complaints in both sectors. Third, just as trustees of private schools can reward administrators who do well, so too can public officials reward administrators of public schools through promotions and other non-monetary benefits. Finally, pride in a job well done may motivates educators and there is no obvious reason why this would differ between sectors.

---

10See Alchian and Kessel (1960).
Sorting

“Sorting” or selection may be as important as incentives in creating a positive and stimulating educational environment. It is essential to ensure that there is a well-qualified teacher applicant pool and that the hiring and retention processes result in selecting the right individuals from the pool. Although teachers are motivated by many factors other than pay, monetary compensation cannot be ignored.

The immediate policy question, then, is “How can we raise teacher quality?” Although money is not everything, money is almost certain to make a difference in this case. Data from 1999 show that teachers on average earn about 77 percent the salary of the average college graduate. This has resulted in a smaller selection of candidates for teaching jobs than would be the case of teaching salaries were higher. Indeed, Caroline Hoxby at Harvard has shown that the average SAT scores for public school teachers are well below the median for the United States as a whole. When the individuals who became teachers were actually applying to college, they themselves were in the lower half of performers on the standard college entrance tests. The result is not particularly surprising, since teachers are so poorly paid relative to other college graduates. It is true, of course, that teachers have more leisure during the summer than do individuals in other occupations, but it is also likely that increasing teachers’ salaries would draw a larger number of people to the profession, and schools could then choose more selectively. In addition, schools would have an easier time replacing those teachers who turned out to be less than effective in the classroom.

Increasing the Supply of Teachers as Compared with Moving Teachers Around:

Some policies are effective in raising the supply of teachers. It is likely that increasing teacher salary
is likely to draw into teaching some college graduates who would have gone into other occupations. Salary increases assist therefore in improving the overall pool of teachers.

Other policies serve primarily to move teachers around, without doing much to the overall pool of teachers. One policy recently suggested by the Governor of California is to give bonuses and forgivable loans to teachers who agree to teach in poor-performing schools. This policy is likely to serve primarily to shift teachers around without doing much to increase the overall supply to the profession. Furthermore, it is hardly clear that moving teachers who were effective in good schools to schools with much poorer performing students will actually improve educational output. Even if it does so for the disadvantaged students, one has to ask what the effects will be on the more able students that they have left behind. The costs to the students in the high achievement schools may be greater than the benefit to those in the low achievement schools. There is some reason to believe that this is so. Most allocations in competitive markets sort positively. The best workers tend to work in the best firms.\footnote{See Abowd, Kramarz and Margolis (1999).} High wage, educated men marry high wage, educated women. In colleges, the best-known professors teach the most able students. If this is the outcome in competitive markets, there is reason to believe that positive sorting is socially efficient because competition maximizes social welfare in the absence of externalities. Thus, a policy that moves teachers instead of changing the nature of the pool is unlikely to have positive social value, although it may have some beneficial redistributive effects.

One problem is that the elasticity of supply of teachers to any one school is much higher than the elasticity of supply of teachers to the profession as a whole. As a result, it is easy to institute policies
that look like they are effective because they induce movement from one school to another. The reality is that they may actually be harmful, having little or no effect on changing the overall pool of teachers.

Tenure:

The institution of tenure, granted to most public and many private school teachers, is often criticized as an extreme manifestation of the unresponsiveness of compensation to performance. The reality is that teachers are only somewhat more protected than other workers from layoff. Even in the non-education sector, once an individual has been with the firm for a few years, the probability of a layoff is very low.\textsuperscript{12} This is particularly true in Europe, where employment protection is common. As a result, it is unrealistic to speak of eliminating tenure. Even if the formal institution were abolished, the informal practice of keeping workers on with very high probability once they have attained sufficient service will reduce the ability to eliminate ineffective senior teachers. Since much of the labor market operates with these ex post rigidities, it is unreasonable to expect that they can be eliminated in education, especially where the forces of market competition are much less direct.

There is little doubt that the institution of tenure reduces the ability to sort effectively. But much of the damage could be eliminated by a small change in the process. The main difficulty with tenure is that it is granted too early in many situations, before teachers’ skills have been accurately assessed. A longer probationary period would probably help significantly in selecting those teachers who have strong “positive fixed effects,” i.e., those who are and are likely to continue to be effective teachers. Longer probationary

\textsuperscript{12}An early study documenting the decline in the turnover hazard with tenure is Mincer and Jovanovic (1981).
periods would also help reduce adverse selection, where the low quality teachers want to stay on because they cannot obtain other jobs and the high quality ones leave. Given the inherent pay compression in the system, waiting a longer time to grant tenure would reduce the number of false positive mistakes made and would thereby diminish the effects of adverse selection.

Conclusion

There are two main conclusions that can be drawn from this essay. First, using pay strategically may be as important in teacher selection as it is in providing teacher incentives. From an incentive point of view, choice of the metric is not unambiguous. Further, making teachers accountable to a particular standard has important distributional consequences by focusing teacher attention on particular students in the class. But pay is very important in sorting, primarily by attracting a larger pool of applicants from which choose. Additionally, since other techniques (like class size reduction) have small direct effects on educational output, channeling a significant portion of money toward higher teacher salaries may be the most cost-effective method of improving educational output.

Second, there are some ways to improve the effectiveness of sorting. Short of eliminating tenure altogether, moving to a longer probationary period may go far toward reducing some of the adverse consequences of the tenure system. In particular, since worker talent and characteristics are relatively stable over time, giving a school district a long enough time to evaluate a teacher may have substantial impact on educational output.
References and Selected Related Work


Flyer, Frederic and Sherwin Rosen. “The New Economics of Teachers and Education,”


