

Teacher Performance Pay:
A Review

Michael J. Podgursky
Department of Economics
University of Missouri – Columbia

Matthew G. Springer
Department of Leadership, Policy, and Organizations
Peabody College of Vanderbilt University

October 24, 2006

NATIONAL CENTER ON
Performance Incentives

This working paper was supported by the National Center on Performance Incentives, which is funded by the United States Department of Education's Institute of Education Sciences (R305A06034). The authors wish to thank Samantha Dalton and Brian McInnis for research assistance, and Jay Greene, James Guthrie, and Robert Costrell for comments. The views expressed in this paper do not necessarily reflect those of sponsoring agencies or individuals acknowledged. For more information on the National Center on Performance Incentives, please visit www.performanceincentives.org.

Abstract

In this paper we examine the research literature on teacher performance pay. Evidence clearly suggests an upsurge of interest in many states and school districts, however, expanded use of merit pay has been controversial. We briefly review the history of teacher pay policy in the U.S. and earlier cycles of interest in merit or performance-based pay. We review various critiques of its use in K-12 education and several strands of empirical research that are useful in considering its likely impact. The direct evaluation literature on incentive plans is slender, focused on short-run motivational effects, and highly diverse in terms of methodology, targeted populations, and programs evaluated. Nonetheless, it is fairly consistent in finding positive program effects, although it is not at present sufficiently robust to prescribe how systems should be designed – e.g., optimal size of bonuses, mix of individual versus group incentives. It is sufficiently promising to support more extensive field trials and policy experiments in combination with careful follow-up evaluations. Future evaluations need to pay particular attention to the effect of these programs on the composition of the teaching workforce, since a growing body of research finds substantial variation in teacher effectiveness as measured by student achievement gains.

Introduction

Salary schedules for teachers are a nearly universal feature of American K-12 public school districts. Data from national surveys show that close to 100 percent of traditional public school teachers are employed in school districts that make use of salary schedules in pay setting (Podgursky, 2006). Thus, roughly 3.1 million public school teachers from kindergarten through secondary level are paid largely on the basis of years of experience and education level – two variables weakly correlated, at best, with student outcomes (Hanushek, 2003).

The single salary schedule tradition contrasts with pay determination practices in the majority of professions where performance-related pay programs are commonplace.¹ In medicine, the pay of doctors and nurses varies by specialty. Even within the same hospital or HMO, pay will differ by specialty field. In higher education, there are large differences in pay between faculty members by teaching field. The average starting assistant professor salary in economics is 1.3 times that of the average starting assistant professor salary in English (Ehrenberg, McGraw, and Mrdjenovic, forthcoming). Moreover, within fields pay is generally merit driven.

Pay determination practices vary between K-12 sectors. Examining early vintages of the Schools and Staffing Survey, Ballou and Podgursky (1997) and Ballou (2001) found that private school teachers were much more likely than their traditional public school counterparts to be rewarded for teaching performance, despite the fact that the majority of private schools reported relying on a salary schedule for teacher pay. Indeed, while less than six percent of traditional public schools reported payments “off schedule”, approximately one-third of private schools reported that teachers were rewarded for excellence in teaching.

1. From a survey of 1,681 businesses, Hein (1996) found that 61 percent used variable, performance-related compensation systems.

Pay determination practices in most professional fields is usually market-driven, enabling organizations to match the offer of competitor firms for employees they wish to retain or create an attractive compensation package for professionals they wish to recruit. Even the federal General Schedule (GS) pay system is more flexible and market-based than those found in most traditional public schools. Civil servants advance through the GS not only in 15 grades, but also along 10 pay steps based on merit and experience (Ballou and Podgursky, 1997). Furthermore, the Department of Defense and Department of Homeland Security within the federal government recently began implementing additional performance-related pay programs to improve organizational performance.

NCLB-induced state accountability systems, coupled with the poor relative performance of U.S. students international math and science tests have stimulated significant interest in the design and implementation of performance-related pay policy. Many districts, and even entire states, are exploring performance-related pay in an effort to bolster administrator and teacher productivity and recruit more qualified candidates. These performance-related pay plans come in many different forms, from compensation based on supervisor evaluations and portfolios created by teachers to payments awarded on the basis of student growth. By some journalistic (perhaps exaggerated) estimates at least one-third of the nation's K-12 public school districts appear "poised" to participate in local, state, or federal-initiated performance incentive policies.² Whether truly "poised" or not, it is clear that many states and districts are actively considering the option. Nor is this interest restricted to the U.S.. A number of European and developing nations have begun to experiment with performance pay (Sclafani and Tucker, 2006).

2. Local and state officials are currently exploring and/or implementing performance-related pay programs in states such as Alaska, Florida, Idaho, Iowa, Minnesota, New Mexico, North Carolina, and Texas. Mark Wallace of the Center for Workforce Effectiveness, a consulting firm specializing in compensation, recently noted that nearly three out of every five states in the nation have enacted legislation requiring localities to explore performance pay (Delisio, 2003).

The purpose of this paper is to examine the economic case for performance-related pay in K-12 education system. While we focus on teachers, by far the largest group of employed professionals in K-12 public education, many of the arguments generalize to school administrators as well. Our review begins with a brief history of U.S. teacher compensation policy and then moves to general descriptions of six large-scale performance-related pay programs currently in operation or about to be launched in US schools. We then review theoretical arguments involving performance-related pay policy, paying particular attention to issues such as performance monitoring, team production, the multi-tasking, and input- versus output-based pay systems. We then review several strands of empirical research that have relevance for this debate, including teacher effect studies, direct evaluations of individual and group performance pay schemes, and studies of incentive pay in private schools and charter schools. While the direct evaluation literature is slender, it does provide some important results for policy. We conclude that while the empirical literature is not sufficiently robust to prescribe how systems should be designed – e.g. optimal size of bonuses, mix of individual versus group incentives – it does make a persuasive case for further experiments by districts and states, combined with rigorous, independent evaluations.

A Brief History of U.S. Teacher Pay Policy

1. Room and Board Compensation Model

The emerging transportation system of America in the early 19th century – by river and canal, and eventually rail – enabled communities situated in rural, agrarian-based locations to trade and prosper. Nearly 80 percent of all citizens living in rural areas and half of all working citizens were farmers (Protsik, 1995). Out of this context emerged the one-room schoolhouse education systems of the late-18th and early-19th centuries whose design was influenced by regional variation in the crop production schedules the dependence of farm production on child labor.

In this environment, the *room and board* compensation model developed. In addition to a small stipend, teachers received room and board by rotating their residences weekly in different students' homes (Protsik, 1995). This facilitated not only attraction and retention of teachers in geographically-isolated locations, but it also clearly solved several principal-agent problems, with each family monitoring in a community monitoring a teacher's ability to instill book-learning as well as to foster the appropriate "moral character" in their children.

However, the one-room schoolhouse education system lacked the capacity to deliver the level or variation in human capital demanded by an industrializing and urbanizing economy (Tyack, 1974). Dramatic increases in the number of students seeking schooling caused a simultaneous increase in the demand for teachers per school. The combined effect of these trends spurred the move toward a grade-based system of education and dramatically altered the nature of teacher compensation.

2. Grade-Based Compensation Model

With industrialization in the late 19th and early 20th century, the "new" economy involved a greater "use of science by industry, a proliferation of academic disciplines, a series of critical inventions and their diffusion" (Golden, 2004). Given the intensified demand for greater skill from a better educated labor force, teacher compensation policy too was re-conceptualized.

The *grade-based* compensation model was created in the late-1800s. Similar to the factory production model preoccupying most sectors of the American economy, the grade-based compensation model paid teachers for the level of skill needed to educate a child at their specified point of educational attainment. Since it was believed that elementary age students were easier to educate, and less formal training was required to teach at that level, teachers who instructed children in their early years earned less than secondary level teachers.

While the design of the new grade-based system made pay uniform by grade level within the profession, the system fostered gender and racial inequities. Entry requirements to teach at the secondary level were more accessible to white males. Furthermore, subjective administrator evaluations of teacher merit were integrated into many grade-based compensation models, resulting in gender- and race-based inequities as well as nepotism (Guthrie, Springer, Rolle, and Houck, 2007).³

3. Position-Automatic or Single Salary Schedule

Around the turn of the 20th century, labor leaders like Samuel Gompers pushed management and factory owners for better working conditions and salaries for their employees. Strikes, boycotts, and negotiations carried out by the American Federation of Labor (1886), the Industrial Workers of the World (1905), and the Congress of Industrial Organizations (1938) were extremely influential in promoting egalitarian pay policy (Kerchner, Koppich, and Weeres, 1997). Collective bargaining for teacher compensation and school district contractual agreements created, and in ensuing years helped maintain, the ubiquitous single salary schedule.

The *single salary schedule*, originally called the “position-automatic schedule”, is a system of uniform pay steps that ensures teachers with the same years of experience and education level receive the same salary (Moelhman, 1927). In a typical schedule, rows indicate years of experience and columns indicate the levels of graduate coursework completed or degrees obtained. This system was implemented to create pay equity, professionalism, and employee satisfaction across grade levels, political wards, districts, and disciplines and to displace prior pay systems negotiated between individual teachers and local school boards (Kershaw and McKean, 1962).

3. Similar findings were reported about a performance-related pay for teachers program that was introduced in England in the 1860s. The ‘payment by results’ program was abandoned after 30 years because “teachers taught to the test, were confined to a narrow, boring curriculum, attempted to arrange the school intake, cheated, ignored bright children and drilled and beat the slower ones until they could satisfy the all-powerful inspector” (Chamberlin, Wragg, Haynes, and Wragg, 2002: 32).

Since its inception, the single salary schedule has been a nearly constant feature of the public school compensation scheme. By 1950, for example, 97 percent of all schools had adopted the single salary schedule (Sharpes, 1987). This figure is remarkably similar to contemporary estimates that 96 percent of public school districts use a uniform salary schedule to compensate teachers (Podgursky, 2006). While the single salary schedule has proved to be remarkably persistent, there have been attempts at change.

4. 20th Century Compensation Experiments in Education

Since first implemented in 1921 in Denver, Colorado and Des Moines, Iowa, the single salary schedule has attracted criticism. Most prominent among these critiques is that the schedule standardizes remuneration, depriving public school managers of authority to adjust an individual teacher's pay to reflect both performance and labor market realities. Numerous teacher compensation reform models have been proposed as alternatives, many under the banner of performance-related pay. The two most prominent types of reform programs have been: (1) merit-based pay and (2) knowledge- and skill-based pay.

Merit-Based Pay: Although merit-based pay programs date back to Great Britain in the early-1700s, it was not until release of the A Nation at Risk report in 1983 that a significant number of public school districts in the United States began considering merit-based pay as an alternative or supplement to the single salary schedule. Merit-based pay rewards individual teachers, groups of teachers, or schools on any number of factors, including student performance, classroom observations, and teacher portfolios. More advanced reward systems hinge on student outcomes attributed to a particular teacher or group of teachers rather than on inputs such as skills or knowledge – a critical distinction that is emphasized later in this review. A report released by the Progressive Policy Institute in 2002 classified school-based performance awards as the most common type of merit-based pay programs operational in the American K-12 public education

school system, but noted as well that rewards can be distributed at, or targeted to, specific grade levels (grade-level teacher teams), departmental units, or combinations thereof (Hassel, 2002).

Knowledge- and Skill-Based Pay: Since the 1990s, knowledge- and skill-based pay has garnered significant attention as an alternative strategy for compensating teachers (Odden and Kelley, 1996). This approach, which has some analogues in the private sector (Beer and Cannon, 2004; Heneman and Ledford, 1998), represents a policy compromise between proponents and opponents of performance-related compensation in education. Knowledge- and skill-based pay programs, such as those designed by the Consortium for Policy Research in Education (CPRE) at the University of Wisconsin, reward teachers for acquisition of new skills and knowledge presumably related to better instruction. Salary increases are tied to external evaluators and assessments (e.g., the Praxis III and National Board for Professional Teaching Standards) that gauge the degree to which an individual teacher has reached specified levels of “competency” (Odden and Kelly, 1996). Although proponents argue that these strategically-focused rewards can broaden and deepen teachers’ content knowledge of core teaching areas and facilitate attainment of classroom management and curriculum development skills (Odden and Kelley, 1996), evidence that the training and credentials being rewarded in these systems is slender (Ballou and Podgursky, 2001; Hanushek and Rivkin, 2004).

Current Performance-related Pay Programs

There is growing interest in federal government in performance-related pay in K-12 public education. While we are aware of no systematic compilation of these programs, groups like the Education Commission of the States (ECS), Mathematica Policy Research, Inc., and federally-funded National Center on Performance Incentives (NCPI) have begun tracking teacher and

administrator compensation reforms and issues.⁴ By all accounts, interest in performance-related pay programs is growing, as is the number of programs under development and being implemented. In this section we consider briefly some current U.S. programs.

1. Denver Public Schools' Professional Compensation System for Teachers (ProComp)

In 1999, the Denver Classroom Teachers Association and the Denver Public Schools reached agreement on an alternative teacher pay plan that linked pay to student achievement and professional evaluations. Following refinement of the pilot model by teachers, principals, administrators and community members, the Professional Compensation Systems for Teachers (ProComp) was adopted in spring 2004 by the Board of Education and members of the Denver Classroom Teachers Association (Community Training and Assistance Center, 2004).

ProComp comprises four components that enable teachers to build earnings through 10 elements, or learning opportunities. These components include: (1) knowledge and skills; (2) professional evaluation; (3) market incentives; and (4) student growth. As noted in Table 1, knowledge- and skill-based pay programs in the form of National Board for Professional Teaching Standard certification holds the greatest potential for pecuniary returns; however, student achievement growth and excellence in professional evaluations can generate significant boost in pay. The achievement growth component includes both teacher and school wide growth awards, with a maximum award of approximately \$2,000.

Insert Table 1 Here

ProComp's place in the Denver Public Schools operational structure was further bolstered by a November 2005 ballot initiative in which Denver voters approved a \$25 million mill levy to fund the plan. Now completing the first of nine voter approved years, ProComp has evolved from a four year pilot program in 16 schools into the nation's most widely-known performance-related pay

4. See, for example, Azordegan et al (2005), www.performanceincentives.org, and Glazerman et al (2006).

program. The ProComp approach is clearly based on “knowledge and skill-based” model described above.

2. Texas’ Governor’s Educator Excellence Award Programs

In 2006, Governor Rick Perry and the 79th Texas Legislature crafted the Governor’s Educator Excellence Award Programs (GEEAP), creating the single largest performance-related pay program in the United States public education system. GEEAP consists of three programs: (1) the Governor’s Educator Excellence Award Program Pilot; (2) the Texas Educator Excellence Grants; (3) and a district-level grant yet to be named. By 2008, GEEAP will provide approximately \$330 million per annum to high-performing public schools in Texas.

Governor’s Educator Excellence Awards Program Pilot. This program is funded at \$10 million annually through the 2008 school year. Funds are distributed in the form of non-competitive grants to schools that meet eligibility criteria. Schools must be in the top third of Texas schools in terms of percentage of economically disadvantaged students and either carry a performance rating of Exemplary or Recognized or be in the top quartile on Texas Education Agency’s Comparable Improvement measure.⁵ Individual award amounts vary according to student enrollment and range from \$60,000 to \$220,000 per year.

Seventy-four schools have been funded to date through this program. Schools are required to use 75 percent of these funds, called Part I funds, for direct incentives to full-time classroom teachers. These incentives are based both on improvement in student achievement and on teacher effectiveness in collaborating with colleagues to improve student achievement on the campus. Part II funds, representing 25 percent of the total award, may be spent on: (1) direct incentives to other school employees (including principals) who contribute to improved student achievement; (2)

5. Comparable Improvement (CI) is a measure that calculates how student performance on the TAKS mathematics and reading/English language arts tests has changed (or grown) from one year to the next, and compares the change to that of the 40 schools that are demographically most similar to the target school.

professional development; signing bonuses for teachers in hard-to-staff subjects, as identified by the Commissioner; (3) teacher mentoring and induction programs; stipends for participation in after-school programs; (4) programs to recruit and retain effective teachers; and/or (5) other programs designed to improve teaching and student achievement.

Texas Educator Excellence Grants. This program is state funded at \$100 million per year. Eligibility criteria and requirements are nearly identical to those of the pilot program. However, schools must be in the top half of Texas schools in terms of percentage of economically disadvantaged students. Grant amounts range from \$40,000 to \$295,000 per year. For the 2006-07 school year, 1,163 campuses are eligible for grants. The Texas Educator Excellence Grants program also separates funds into Part I and Part II funds, with the former based on objective measures of student performance and the latter on a variety of incentives and professional growth activities. Part II funds, representing 25 percent of the total award, may be spent on the same activities outlined in the pilot program, with the addition of stipends for teachers certified in-field or holding post baccalaureate degrees.

District-Level Grant. This program will be funded at approximately \$230 million annually with state funds provided through the Texas Educator Excellence Fund. All districts in the state will be eligible for funding. Districts may apply for funds for all campuses or for selected campuses. Districts are required to use at least 60 percent of funds to directly award classroom teachers based on improvements in student achievement. Remaining funds may be used: (1) as stipends for mentors or teacher coaches, teachers certified in hard-to-staff subjects (as identified by the Commissioner), teachers certified in their field, or teachers who hold post baccalaureate degrees; (2) as awards to principals based on improvements in student achievement; or (3) to implement components of Milken Family Foundation's Teacher Advancement Program.

3. Florida's Special Teachers Are Rewarded (STAR)

The 2006-07 budget approved by the Florida State Legislature included a \$147.5 million appropriation within the Florida Education Finance Program (FEFP) for the Special Teachers Are Rewarded (STAR) performance-related pay program. Suspending the 2001 State Board of Education Performance Pay Rule, known as E-Comp, Florida's new STAR program requires that all traditional public schools and public charter schools integrate a performance-related pay program into the existing salary schedule. Districts are requested to submit a STAR implementation plan prior to January 2007 with the expectation of approval in April 2007.

STAR has four major components: (1) eligibility declaration; (2) determination of number of rewards; (3) evaluation instrument; and (4) instructional personnel evaluation based on student performance. Guidelines for the first two components require that all instructional personnel be eligible for a STAR award, and that the majority of instructional personnel are rewarded similarly. Indeed, each district must design the allocation mechanism to award a minimum of 25 percent of instructional personnel with fiscal rewards at or above 5 percent of their current base salary. The third and fourth components – evaluation instrument and personnel evaluation based on student performance – require each district to develop criteria for assessing academic improvement as well as methodology for monitoring students' progress both within individual courses and within scholarly disciplines throughout various segments of their academic career. After these provisions are met, the remaining STAR funding may be used to further award excellence among faculty.

4. Minnesota's Q-Comp

In July 2005, the Minnesota State Legislature approved Q-Comp, a performance-related pay program for teachers. Q-Comp incorporates both traditional career ladders and professional development for teachers, while advancing existing state standards through integration of measures to compensate teachers according to state approved measures of student achievement. Under Q-Comp guidelines, 60 percent of any compensation increase must be based on district professional

standards and on classroom-level student achievement gains. Q-Comp presently operates in only 22 of 348 regular school districts across the state, however in the next two years 134 school districts have indicated intent to submit a Q-Comp proposal to the state. District plans that are approved by the state department of education can be awarded up to \$260 more per student to support implementation and sustenance of their merit-based compensation plan.

5. Milken Family Foundation's Teacher Advancement Program (TAP)

The Teacher Advancement Program (TAP) was developed in 1999 by the Milken Family Foundation, a philanthropic organization based in Santa Monica, California, to increase the number of highly qualified teachers, improve instructional effectiveness, and enhance student achievement.⁶ TAP consists of four major components: (1) multiple career paths; (2) on-going applied professional growth; (3) instructionally focused accountability; and (4) performance-related compensation.

Multiple Career Paths. TAP's multiple career paths position high quality teachers to pursue a variety of positions, advance professionally, and earn higher salaries without having to abandon the classroom. If teachers demonstrate consistent success they have the opportunity to become career, master, or mentor teachers. This option of multiple career paths is of particular importance considering that the current state of career advancement in the United States education system is structured to remove experienced and qualified teachers from the classroom. That is, added pay, recognition and prestige are gained by moving into administrative positions, thereby sending a subtle message to teachers that career advancement occurs outside the classroom (Ballou and Podgursky, 1997; Guthrie and Springer, 2006).

Ongoing Applied Professional Growth. TAP allocates time during the instructional day for teachers to meet and collaborate on instructional and curricular issues. These meetings are either group or individual-focused and often scheduled with a TAP-identified mentor or master teacher

6. TAP was recently renamed the National Institute for Excellence in Teaching (NIET).

within a particular school or district. TAP's mission of ongoing applied professional growth provides a framework for teachers to: (1) set learning goals based on analysis of students' performance; (2) identify proven research-based learning strategies to address goals; (3) work collaboratively to develop new instructional practices; (4) bring these practices to the classroom; and (5) measure how well newly-formulated strategies help students meet learning goals.

Instructionally-focused Accountability. Instructionally-focused accountability refers to TAP's mechanism for evaluating teachers. In an effort to assess teacher performance appropriately, TAP employs a grading rubric to measure systematically a teacher's content knowledge, instructional methods, and student learning gains. These evaluations are ultimately used to determine a teacher's career ladder advancement within the school.

Performance-related Compensation. TAP's construct of performance-related compensation rewards teachers across three dimensions: (1) student performance; (2) increased roles and responsibilities; and (3) classroom teaching performance. In linking pay to these three dimensions, TAP's remuneration mechanism represents a substantial departure from more traditional practices in which teacher pay is based on years of experience and highest degree held.

TAP currently operates in more than 125 schools in 9 states and 50 districts. Another 10 states presently are pursuing program implementation in routinely low-performing schools. In the aggregate, there are approximately 3,500 teachers and 56,000 students in TAP schools across the country.

6. United States Department of Education's Teacher Incentive Fund

Congress recently appropriated \$99 million per annum to school districts, charter schools, and states on a competitive basis to fund development and implementation of principal and teacher performance-related pay programs. As part of the United States Department of Education's fiscal year 2006 Appropriations Act (P.L. 109-149), the Teacher Incentive Fund (TIF) is a direct

discretionary Federal grant program. Reports indicate that TIF will fund approximately 10 to 12 performance-related compensation projects with an estimated per-project award size of \$8 million per year. Awards are expected to be announced in October 2006 with a January 2007 start date.

Theoretical Arguments For and Against Performance-related Pay Programs

As noted above, in the wake of the A Nation at Risk report in 1983, a number of school districts experimented with performance-related pay programs as a means to improve student outcomes and reform the single salary schedule. Research on these programs highlighted the difficulty inherent in creating a reliable process for identifying effective teachers, measuring a teacher's value-added contribution, eliminating unprofessional preferential treatment during evaluation processes, and standardizing assessment systems across schools (e.g., Murnane and Cohen, 1986; Hatry, Greiner, and Ashford, 1994).⁷ Criticisms stemming from these generally short-lived programs have since stigmatized more recent attempts to devise and implement performance-related pay programs.

Murnane and Cohen (1986) is one of the more influential critiques of this early wave of merit-based pay programs. Drawing on personnel economics literature, they argued that post-1983 merit-based pay plans failed because teaching is not a field that lends itself to performance-related compensation, a perspective that Goldhaber, et. al. (2005) recently termed the “nature of teaching” hypothesis. Given its influence, and the fact that subsequent critiques have often raised the same arguments, we devote some attention this article.

1. The “Nature of Teaching” Hypothesis

7. Past programs have also failed due to insignificant financial incentives for successful teachers, teacher unions who were opposed to alternative compensation systems, and lack of an evaluation process that could assess outcomes and recalibrate programmatic components to bring the program to scale (see, for example, Ballou and Podgursky, 1997; Guthrie and Springer, 2006).

Performance Monitoring. A major argument against merit-based pay programs concerns the difficulty in monitoring teacher performance. According to Murnane and Cohen, teacher performance is more difficult to monitor than performance in many other professions because output is not readily measured in a reliable and valid manner.⁸ Unlike, say, the sales of a salesman or the billable hours of a doctor or lawyer, the output of a teacher is not marketed. Thus, the education sector cannot readily measure the value of the services provided by an individual teacher or group of teachers.

While this argument no doubt had merit at the time, its relevance may be waning given the major advances in data systems being put in place in states and districts. States and districts are rapidly developing massive longitudinal student-level databases that permit for more precise estimation of value-added contributions at the building-, grade-, and, in a growing number of states, teacher-level. Furthermore, the United States Department of Education has also created a competitive grant programs to encourage states to develop longitudinal data systems that support value-added measurement. As data and measurement systems grow in sophistication the measurement of teacher and school performance will also become considerably more reliable and valid.⁹

8. Similar points were made following implementation of the single salary schedule in 1921. In one of the first education finance textbooks published, Moehlman (1927) argued for development of a salary schedule that provides “as scientifically possible for the best returns to society for the increasing public investment” by approaching salaries from “its economic and social aspects and not in terms of sentimentality”. However, he concluded that an objective and standardized system for determining merit did not exist nor was there capacity to develop a school or district-level system. Consequently, the most relied upon method for evaluation of merit was a single salary schedule called the “position-automatic schedule” which automatically advanced teachers by annual pay increments ranging between \$50 and \$200 after their first year of teaching until a predetermined maximum salary was reached somewhere between 10-15 years of service.

9. See Guthrie (2006) for a detailed account of data deficiencies and needs in modern educational research, practice, and policy.

In spite of these technological advances, to the extent that these new performance-related pay programs rely on estimates of teacher value-added, it is important to note that there are still concerns about the statistical reliability and robustness of these value-added estimates. We do not know how successfully these statically-determined estimates of teacher performance effects can guide educational practice and be harnessed as incentives for teachers to change practice. Indeed, researchers express caution in interpreting teacher effects purely as an attribute of the teacher without consideration of the school context and with the stability of these measures over time (McCaffrey et al, 2003; McCaffrey et al, 2004; Ballou, Sanders, and Wright, 2004; Ballou, 2005; Koedel and Betts, 2005).¹⁰

Team Production. A second argument against merit-based pay programs concerns team production. To considerable extent, teachers work as members of a team. Introducing performance-related rewards at the individual teacher level might reduce incentives for teachers to cooperate and, as a consequence, reduce rather than increase school performance. Some scholarship argues that the team dynamic can be destroyed between teachers as well as between teachers and administrators, especially if administrators are put in a position of rewarding individual teacher performance.¹¹ Of course, this is a criticism of individual performance-related pay programs. A performance bonus given to an entire team of teachers would not undermine team morale. This is especially germane considering most teachers work in relatively small teams, and economic literature suggests team incentives may work quite well in small teams because there is mutual monitoring coupled with an

10. As noted by Beer and Cannon (2004), business compensation literature has argued that chances for successful implementation of performance-related compensation systems might be increased in organizations where (1) the culture discourages opportunism, (2) top management reinforces this culture by its example, and (3) employees have long-term careers or professions in which their reputation is a valuable commodity.

11. A similar argument was explored in response to a performance-related pay scheme that was introduced in England and Wales during the 2000-2001 school year. Adnett's (2002: 145) economic analysis of what he termed, "threat to the collegiate ethos", suggested that, "the presence of asymmetric information, externalities, and teamwork effects can provide a rationale for encouraging professional motivation."

easy information flow among teams members and options for subjects to reciprocate among each other within the team (Kandel and Lazear, 1992; Vyrastekova, Onderstal, and Koning, 2006).¹²

The Multi-Tasking Problem. Another theoretical criticism of performance-related pay programs in the literature concerns the issue of multi-tasking when relying on tests or other quantitative measures of teacher performance (e.g., Holmstrom and Migrom, 1991; Hannaway, 1992; Dixit, 2002). This problem arises when the performance of a worker has multiple dimensions, only some of which are measured and incentivized. When there is structural misalignment between an organization's overall mission and the activity to which incentives are attached, not surprisingly, employees tend to shift work toward the metered, rewarded activity, and away from other important activities.

An important concern in this regard is “teaching to the test” – an education catch-phrase used to describe narrowing of curriculum in an effort to elevate student test scores. Teachers' contributions to student learning are multifaceted; however, if an inordinate amount of weight is placed on student assessments then other valuable activities might be slighted. In the general personnel literature, the solution to the multi-tasking problem is to diversify the measures used to evaluate performance, such as supervisor evaluations or other broad-based assessments to complement quantitative measures.

Incentive schemes that tie teacher pay to achievement gains by students – whether at the individual teacher or “team” level – may create more opportunities for cheating or other opportunistic behavior in the long run. For example, studies of high stakes accountability systems have documented teachers focusing excessively on a single test and educators altering test scores and/or assisting students with test questions (Goodnough, 1999; Koretz et al, 1999; Jacob

12. Of course, rewards to an entire team, rather than to each member, introduce the “free rider” problem. If a team member exerts effort and raises overall team output by X, he will receive only receive a return of X/N , where N is the size of the team. Clearly as N grows the performance incentive shrinks rapidly. For a discussion of this problem, see Prendergast (1999).

and Levitt, 2005). Related analyses have found evidence of schools' strategic classification of students as special education and limited English proficiency (Deere and Strayer, 2001; Figlio and Getzler, 2002; Cullen and Reback, 2006; Jacob, 2005), use of discipline procedures to ensure that low-performing students will be absent on test day (Figlio, 2003), manipulation of grade retention policies (Haney, 2000; Jacob, 2005), misreporting of administrative data (Peabody and Markley, 2003), and planning of nutrition enriched lunch menus prior to test day (Figlio and Winicki, 2005). Together, these findings suggest the need for multiple mechanisms in pay for performance programs to minimize negative spillover effects within high-stakes contexts.¹³

2. Payment for Input and Payment for Output

Edward Lazear, a major contributor to the “new personnel economics” literature, provides a useful conceptualization of the performance-related pay problem in K-12 education, and assesses the economics of alternative teacher compensation regimes which he terms payment for input and payment for output (Lazear, 2003). In the absence of externalities or information problems, payment for output always trumps payment for input in terms of raising overall productivity. Two principals reasons – hiring practices and labor market selection – are discussed below.

Hiring Practices. District and building administrators are plagued by an informational deficiency when hiring teachers and other instructional staff. This necessitates that principals use noisy signals of true teacher effectiveness (e.g., years of experience, highest-degree held, past-employer recommendations).¹⁴ Informational deficiencies in the hiring process are ameliorated in most professions by subsequent employee performance assessments, and as pay raises become more

13. Dixit (2002: 719) provides a laconic assessment of the evaluation problem, “To sum up, the system of public school education is a multitask, multiprincipal, multiperiod, near-monopoly organization with vague and poorly observable goals.”

14. Input-based compensation programs such as knowledge- and skill-based pay rely upon a series of noisy signaling indicators (e.g. portfolios, educational artifacts, few evaluations) similar to those in the existing single-salary schedule, and thus are not a suitable alternative to existing labor market inefficiency.

closely tied to actual productivity, thereby lessening dependence on input-based indicators for employees (Altonji and Pierret, 1996). Of course, the single salary schedule, along with teacher tenure, makes it difficult for pay and performance to align after hire. For example, if only effective teachers have their contracts renewed then pay on the basis of seniority would tend align pay and performance. While such a mechanism may work in the first probationary years of teacher employment, after teachers earn tenure, contract non-renewal can only be triggered by severe malfeasance on the part of the employee.

Labor Market Selection. Lazear also discerned a more subtle, but important factor in the gains from a performance-related, or output-oriented, pay system that arise from labor market selection. A performance-related pay program will tend to attract and retain individuals who are particularly good at the activity to which incentives are attached, and repel those who are not. He noted that this effect on the workforce can be very important in explaining productivity gains. For instance, in one of his own case studies outside of teaching, Lazear (2000) found that sorting effects were both substantial and roughly equal in magnitude to motivation effects. In other words, while the incentive system raised the productivity of the typical worker employed, it also tended to raise the overall quality of the workforce.

Some researchers speculate that this selection effect may be a significant factor in teacher labor markets. For example, studies of teacher turnover consistently find that high ability teachers are more likely to leave teaching than teachers of lower ability (Murnane and Olsen, 1990; Podgursky, Monroe, and Watson, 2004). This trend may be due to constraints on wages rather than the attraction of other market opportunities. A recent provocative study by Hoxby and Leigh (2004) found evidence that the migration of high ability women out of teaching between 1960 to the present was primarily the result of the “push” of teacher pay compression – which took away relatively higher earnings opportunities for teachers – as opposed to the pull of greater non-teaching

opportunities. Although the remunerative opportunities for teachers of high and low ability grew outside of teaching, it was pay compression within the education system that accelerated the exit of higher ability teachers. To the extent that these high ability teachers were more effective in the classroom, a performance-related pay program likely would have kept more of them in teaching.

Lazear's selection arguments also undermine one other critique of teacher merit pay by Murnane and Cohen. These authors argue that in any effective merit pay system employers should be able to tell workers what they need to do in order to become more effective. In other words, if ineffective teachers do not know what to do in order to raise their performance, and supervisors cannot provide such guidance, then the motivational effect of merit pay will be nil. However, if the underlying range of teacher effectiveness is great (and evidence considered below suggests that this is the case), then simply tying pay to performance may significantly raise performance even if no individual teacher's productivity rises, simply through differential recruitment and retention of high performance / high paid teachers.

Empirical Research

Economic theory can take us only so far in predicting the effect of teacher performance pay. Ultimately we must turn to the data. In this section we review four strands of research relevant to the debate: (1) the teacher effects literature; (2) studies linking teacher effects to performance assessments; and (3) direct evaluations of individual- and group-based performance pay programs, and (4) evidence from private and charter schools

1. Teacher Effects Studies

Over the last decade, researchers have begun to exploit massive longitudinal student achievement data files to undertake "value-added" studies of teacher effectiveness. Beginning with William Sanders' work in developing the Tennessee's Value Added and Assessment System

(Wright, Horn, and Sanders, 1997; Ballou, Sanders, and Wright, 2004), teacher value-added studies have expanded to states such as Texas (Rivkin, Hanushek, and Kain, 2005) and to large school districts such as New York City Public Schools (Kane, Rockoff, and Staiger, 2004; Boyd, Grossman, Lankford, and Loeb, 2006) and Chicago Public Schools (Aronson, Barrow, and Sanders, 2003). These studies have consistently found evidence of large and somewhat persistent differences in achievement gain scores between classrooms and teachers, suggesting that teachers can have a substantial effect on student achievement growth, particularly if teacher effects are cumulated over a number of years.

While researchers have found substantial variation in teacher effects within school districts, and even within schools, they also have consistently found that these effects are largely unrelated to measured teacher characteristics such as the type of teaching certificate held by the teacher, a teacher's level of education, licensing exam scores, and experience beyond the first two years of teaching. Indeed, nearly every researcher conducting rigorous teacher effect studies has taken note of this fact (e.g., Kane, Rockoff, and Staiger, 2006; Rivkin, Hanushek, and Kain, 2005; Aronson, Barrow, and Sanders, 2003; Goldhaber and Brewer, 1997). For example, in a large scale study of certification status and new teacher effectiveness in New York City Public Schools, Kane, Rockoff, and Staiger (2006: 40) write:

...there is not much difference between certified, uncertified, and alternately certified teachers overall, but effectiveness varies substantially among each group of teachers. To put it simply, teachers vary considerably in the extent to which they promote student learning, but whether a teacher is certified or not is largely irrelevant to predicting their effectiveness.

We have reproduced from their study a chart of estimated teacher effects that demonstrates clearly this point. Figure 1 reports variation in estimated teacher effects for new teachers by type of

teaching certificate held.¹⁵ It is readily apparent that the distributions overlap almost entirely, illustrating negligible differences between certified, uncertified, and alternately certified teachers. With respect to merit-based pay, note the very wide variation in teaching effectiveness within each certification group. Any policy that can retain and sustain the performance of teachers in the upper-tail of the distribution, and enhance the performance of or counsel out teachers in the lower-tail, possesses potential for substantial impact on student growth.

Insert Figure 1 Here

A recent study of Chicago Public School teachers by Aaronson, Barrow, and Sanders (2003) further illustrates this point. Like other such studies, this work was based on a very large longitudinal file of student achievement scores linked to teachers. What makes this study unique is that the authors had very extensive administrative data on teacher characteristics heretofore unavailable in other studies, including education, experience, types of teaching licenses, and selectivity of the teacher's undergraduate college. Aaronson and colleagues found that over 90 percent of teacher effects are not explained by any of these measured teacher characteristics.

The fact that most studies to date conclude that teacher graduate degrees – the most common educational credential – have a marginal effect at best on student achievement (Hanushek, 2003), there is little empirical support for the current credential-based teacher compensation system. The National Board for Professional Teaching Standards Certification has been promoted as an alternative to merit-based pay programs (National Commission on Teaching and America's Future, 1996). However, even here the evidence concerning performance is mixed (Goldhaber and Anthony, 2006; Sanders, Ashton, and Wright, 2005).¹⁶

15. Another team examining New York City Public Schools reached a similar finding (Boyd et.al., 2006).

16. Angrist and Guryan (2005) estimate the effect of state teacher testing requirements on teacher wages and teacher quality as measured by educational background and conclude that state-mandated teacher testing increases teacher wages with no corresponding increase in quality.

The widely-dispersed and idiosyncratic nature of teacher effects has important consequences for the performance pay debate. On the one hand, it suggests that credential-based pay reforms are not likely to have substantial effects on student achievement. On the other hand, it points to substantial student achievement gains if the mix of low and high performing teachers can be altered. A policy that ties pay to performance over time would likely recruit or retain more teachers in the upper tail of ability into the teaching workforce, and encourage low productivity teachers to either improve or leave for non-teaching positions. For example, suppose that a pay system is monotonically related to the teacher effectiveness measured on the horizontal axis in Figure 1. Further assume that all teachers have identical non-teaching earnings, illustrated by a vertical bar somewhere on the horizontal axis indicating the teaching productivity equivalent of the alternative wage. Ignoring, for the moment, non-pecuniary preferences for teaching versus other jobs, teachers with productivity to the right of the vertical bar would move to, or stay in, teaching, and those to the left would exit. Teacher turnover would thus become part of a virtuous cycle of quality improvement, rather than a problem to be minimized.

2. Teacher Effects, Performance-related Awards, and Principal Evaluations

The multi-tasking problem identified in previous critiques of performance-related pay programs makes the case for subjective assessment by supervisors as part of a multi-factor assessment. The assumption is that the supervisor evaluation picks up important teacher behaviors that student achievement gains do not. Glewwe, Ilias, and Kremer (2003), discussed in more detail below, raise this issue in the context of their assessment of a short-lived teacher incentive scheme in Kenya. However, independent of the multitasking issue, it is useful to know the strength of association between supervisor evaluations and student achievement gains where this relationship can be measured. This at least increases our confidence in supervisor measures in contexts in which they cannot be validated by test score gains (e.g., music or social studies teachers). A small number

of studies have examined the relationship between these subjective assessments and teacher performance as determined by student test score gains. Indeed, and as early as the mid-1970s, a number of educational researchers have concluded that principal evaluations are a reliable guide to identifying high- and low-performing teachers as measured by student test-score gains.

Two older studies using student longitudinal data are relevant (Armor et al., 1976 and Murnane, 1975). These studies found large effects of principal evaluations on student achievement gains. More recently, Sanders and Horn (1994: 2000) reported that, “There is a very strong correlation between teacher effects as determined by the data and subjective evaluations by supervisors.”

In a particularly rigorous study focused entirely on the predictive validity of supervisor evaluations, Jacob and Lefgren (2005) assessed the relationship between teacher-performance ratings as identified on a detailed principal evaluation and teacher effects as measured by student achievement gains. In estimating teacher effectiveness measures for 202 teachers in grades 2 through 6 in math and reading, Jacob and Lefgren found a statistically significant and positive relationship between value-added measures of teacher productivity and principals’ evaluations of teacher performance.¹⁷

Another interesting dimension of this study was an “out of sample” prediction of 2003 student achievement scores based on principal ratings and teacher value-added estimates from 1998 through 2002. Students had higher average scores in math and science if they had teachers with not only higher measured teacher effectiveness in prior years, but also higher principal ratings. Jacob and Lefgren found further that the principal evaluation remained a statistically significant predictor of current student achievement even when teacher value-added (in the previous year) was added to the model. This finding suggests that principal evaluations provide an important independent source of

17. Among other things, principals were asked to assess the ability of teachers to raise student achievement on a scale from 1 (inadequate) to 10 (exceptional).

information on teacher productivity.

Although these studies tend to indicate that principals are relatively adept at identifying above- and below-average teachers, it is important to question whether subjective assessment practices persist in a performance-related pay system. The fact that a principal identifies a teacher as “inadequate” on an anonymous survey does not mean necessarily that she will do so in a high-stakes environment. Indeed, a primary reason the single salary schedule replaced the grade-based compensation system was that subjective measures used to reward teachers were highly susceptible to gender and racial discrimination as well as nepotism.¹⁸

Two studies shed some light on whether “old style” merit plans that were based in part on supervisor evaluations are positively associated with measures of teacher productivity. Cooper and Cohen (1997) found that classroom gain score measures were higher for teachers who received merit pay awards in South Carolina. However, in the individual pay component of the plan, teachers who applied for the award were evaluated on four criteria, one of which was a performance evaluation and another was evidence of superior student achievement gains. Accordingly, these findings are not considered a strong test of the hypothesis.

A more recent study by Dee and Keyes (2004) examined the relationship between career ladder bonuses and student achievement gains in Tennessee’s Project STAR data. What makes this study unique is that students were randomly assigned to teachers in the experiment. While the focus of the STAR experiment, and subsequent research studies, has been the effect of class size, Dee and Keyes take advantage of the fact that students were also randomly assigned to Tennessee career ladder teachers. Teachers advanced on the career ladder rungs primarily on the basis of subjective evaluations typically conducted by a local principal. Dee and Keyes found that teachers with career

18. Marsden and Belfield (2006) report results from a panel survey of classroom and head teachers’ views of a performance-related pay system in England and Wales. They found the number of teachers who believed managers would use subjective evaluations to reward favorites dropped from over 50 percent when the program was first introduced to less than 20 percent four years later.

ladder status (i.e., those who have passed one or more evaluations) were more effective than teachers who had not obtained career ladder status. However, it is not possible to sort out the selection effect from the validity of the evaluation per se. It may have been the case that teachers who self-select to join the program or advance up the ranks are simply better teachers.

While no single study is definitive in this area, a small literature has developed showing that principal evaluations and performance-related promotions and/or awards based in whole or in part on principal evaluations are associated with higher classroom teacher effectiveness as measured by student achievement gains. This finding does not entirely address the multi-tasking problem that recognizes the presence of many valuable attributes to teacher performance not adequately measured by state assessments. Nonetheless, to the extent that principals' subjective assessments capture these attributes, it is useful to know that these evaluations are also correlated with teacher productivity when measured by student gain scores.

3. Assessments of Performance-related Pay Programs

While there have been numerous experiments in individual and group incentive pay for teachers over the years, the evaluation literature is very slender. Table 2 lists studies that we have found in the literature that employ a conventional treatment and control evaluation design. Interestingly, in contrast to the very mixed findings of studies of teacher characteristics such as certification or teacher education (Hanushek, 2003), the incentive literature generally finds positive achievement effects. Table 2 summarizes some characteristics of these performance-related pay programs and findings from relevant evaluation studies.

Insert Table 2 Here

It should be noted that we have not attempted a more sophisticated “meta-analysis” or analytical synthesis for several reasons. First, these programs involve very different incentive schemes or “treatments.” Second, the outcome variables analyzed also vary considerably, sufficiently

so that we do not feel it is useful convert them to a common metric. Third, the number of studies is small, the range of assessments considered is highly diverse, and the evaluators did not always have access to the outcome variable of interest. In some cases it is clear the performance-related pay programs were poorly designed, and incentives were not necessarily aligned with the educational outcome of interest to a school and/or district. A. The last column represents our assessment of the outcome of the study. Although it is our subjective assessment, we do not believe that the authors, or most other careful readers, would dispute these interpretations. Since the number of studies is small, and the range of assessments considered is highly diverse, we will briefly discuss each.

Ladd (1999) and Clotfelter and Ladd (1996) examined the effect of a school-wide incentive scheme implemented in the Dallas Independent School District in the mid-1990's. The Dallas Accountability and Incentive Program provided a modest pay boost to all teachers in high-performing schools. Since the program was intended to raise the performance of all schools in the district, the district was the appropriate treatment unit in the authors' analyses. While data aggregated to the district level certainly make program evaluations of this nature challenging, the authors found that achievement in Dallas rose relative to other Texas public school districts.

All other studies identified in Table 2 focused on school or teacher level effects. Two of the most methodologically rigorous evaluations were conducted by Lavy (2002; 2004). In both of these studies the performance-related program was a fixed-tournament designed to raise pass rates on high school exit exams in low socioeconomic high schools in Israel. Although schools were not randomly assigned to a control or treatment condition, both programs were implemented in a way that permitted rigorous non-experimental evaluation. Furthermore, the incentive schemes were carefully designed so as to minimize gaming or other opportunistic behavior.

The first study considered a tournament in which a selected group of low-performing schools competed on the basis of school-wide performance (Lavy, 2002). Bonuses ranged from \$200

to \$715, and were distributed equally to all teachers in the winning schools. Results indicated a positive effect on participating schools relative to non-participating control schools.

The second study examined an individual teacher bonus program, also run as a tournament (Lavy, 2004). Essentially, teacher participants were ranked on the basis of value-added contributions to student achievement on a variety of exit exams, and bonuses were given to top performing teachers. The bonuses were substantial; as large as \$7,500 per class on an average base pay of \$25,000. Results indicated a positive effect in that participating teachers (i.e., both bonus recipients and non-recipients) performance rose relative to that of a control group that did not participate in the incentive program.

Lavy (2004) also investigated whether the program exhibited the type of negative spillover consequences often discussed in the “contracting” literature. First, as proposed in the multi-tasking problem, test scores in other non-tournament subjects did not fall. Moreover, and consistent with the teacher value-added literature discussed above, teacher characteristics such as experience or certification could not predict the winners. Another attractive feature of this study is that Lavy compared the cost effectiveness of the individual bonus scheme with that of group bonuses or policies providing additional educational resources, save for pay, to low achieving schools. He found that the cost per unit gain in the incentive program dominated those found in the group incentive or added resource programs.

Interestingly, even in those cases where Table 2 reports “mixed” findings, the incentive scheme always had a positive effect on the behavior to which incentives were attached (i.e., “You get what you pay for”). Eberts, Hollenbeck, and Stone (2002) studied the effect of an incentive scheme in a single alternative high school in Michigan. In response to a growing dropout rate problem, the school introduced a bonus system that paid teachers to raise their students’ course completion rates. The researchers compared the “treatment” school to another alternative high school considered

comparable. The bonus program significantly raised course completion, but, not surprisingly, non-targeted variables such as student pass rates or grade point average dropped because academically marginal students were induced to stay in school. Clearly a better performance pay plan would have incorporated a larger set of performance indicators. However, the results of the study show that teachers responded to a short-term incentive plan, and raised the course completion rate.

A second “mixed” study involved primary schools in rural Kenya. Glewwe, Ilias, and Kremer (2003) study is unique in its use of random assignment methods whereby fifty schools were chosen at random for participation from among one hundred relatively low-performing rural schools. Tied to student pass rates on state exams in a variety of subject areas, the allocated bonuses were substantial, ranging from 21 to 43 percent of monthly pay. However, the program was of limited duration (originally announced as one year, later extended to two). Glewwe and colleagues found increased pass rates on the state exams, the target of the incentive, but those gains did not persist in subsequent years, which they took as evidence of gaming on the part of teachers. While targeted teachers provided more after school test preparation, the researchers found no evidence of differences in pedagogy or in teacher absenteeism, traditionally a major problem in the schools under study.

Both the Michigan and Kenya cases thus showed that teachers responded to the incentives in the program, but that the incentive regimes were poorly designed. A review of the principal-agent multi-tasking literature outside of education by Courty and Marschke (2003) highlights the dynamic learning context of these incentive systems. Their model illuminates the importance of experimentation and trial and error in principals’ efforts to develop a reliable performance measure. In this respect, these two “mixed” studies, and Courty and Marschke’s review, point to the need for experimentation and rigorous evaluation of performance-related pay programs before dismissing their potential utility.

Atkinson et al (2004) evaluated the effect of a teacher bonus pay scheme in the United Kingdom. The program was short-lived – only one year in duration. In addition, it turned out ex post that the bonus was provided to nearly all teachers who applied. Atkinson and colleagues, however, made a fairly convincing case that this outcome was not known to teachers in advance. To win the performance-related award, teachers were required to provide evidence to the United Kingdom’s Education Ministry that the achievement gains of their students exceeded national averages in five areas. Atkinson et al compared gain score data for eligible teachers with benchmark data on gain scores prior to the implementation of the program. It is important to note, however, that the authors had some difficulty in developing a representative sample with pre- and post-program gain-score data; nonetheless, they did find a large and statistically significant effect of this one year program.

Figlio and Kenney (2006) analyze a sample broadly representative of American K-12 schools. By creatively merging data from the National Educational Longitudinal Survey of 1988, their own survey on merit pay, and the 1993-94 Schools and Staffing Surveys, they were able to exploit the natural variation in the use of incentive-based pay among both public and private schools.¹⁹ While other studies analyzed only the presence or absence of a given program, the natural variation in incentive programs in Figlio and Kenney’s sample enabled construction of a school-level measure of the strength of the teacher incentive “dosage” reflecting not only the existence of a merit-based pay scheme but also its pecuniary consequences. The effects of even modest doses of incentive pay were statistically significant in both public and private schools. Figlio and Kenney also conclude that the effect of a high level of implementation of incentives relative to none at all

19. Clearly using natural variation has its cost, since the variation may not arise exogenously. However, one benefit of a study using natural variation is that many of the schools using the incentive plans may have had them in place for a sufficient length of time to pick up both motivation and selection effects.

impacts achievement comparable to a one standard deviation increase in days absent for the average student and an increase in maternal education of three years.

4. Incentive Pay in Private and Charter Schools

If contracting problems in K-12 public education such as performance monitoring, multi-tasking and team production are inherent in the production process, and sufficiently severe so as to preclude group or individual performance-related pay programs, then we would expect to see similar pay structures in charter schools and private schools when compared to traditional public schools. Several studies have examined this question and found significant differences.

Hoxby (2002) hypothesizes that increased competition leads to greater use of merit and performance-related pay and finds evidence in support of that thesis. Ballou (2001) directly explores the question of whether private schools make greater use of performance-related pay by analyzing the structure of earnings conditional on experience and education. His findings point to greater use of merit pay in private school wage setting. A more recent study by Podgursky (2006) examines data from the 1999-2000 Schools and Staffing Surveys (see Table 3). He finds significantly greater use of performance-related pay bonuses by private and charter schools. Finally, Ballou and Podgursky (1993) examine survey data on teacher attitudes from the 1987-88 Schools and Staffing Surveys. They find that private school teachers are much more supportive of performance-related pay than public school teachers, which is consistent with the sorting hypothesized by Lazear (2003).

Insert Table 3 Here

Conclusion

In this paper we examine the economic case for performance-related pay in K-12 education. Our focus is on teachers, by far the largest group of employed professionals. However, many of the arguments generalize to school administrators as well. We began with a brief history of teacher

compensation policy from the 18th century to present and then moved to a description of six large performance-related pay programs currently operating in the American K-12 public education system.

We also reviewed several ideas from the growing personnel economics literature that have particular relevance for teacher performance pay. We considered some well-known problems in the use of performance-related pay programs in any organizational context and their relevance for K-12 education. One important theme, which is not widely considered in the education studies, is motivation versus selection effect in an incentive system. A second important theme that emerges is the role of credentials versus performance in pay determination.

Prior studies of performance-related compensation were extremely diverse in terms of incentive design, population, type of incentive (group versus individual), strength of study design, and duration of the incentive program. While the literature is not sufficiently robust to prescribe how systems should be designed – e.g., optimal size of bonuses, mix of individual versus group incentives – it is sufficiently positive to suggest that further experiments and pilot programs by districts and states are in order. As these programs are introduced, however, it is critical that they be introduced in a manner amenable to effective evaluation. Moreover, as noted by Courty and Marschke (2003), an overarching lesson seems to be that trial and error is likely required to formulate the right set of performance incentives. Development of massive student longitudinal achievement databases opens prospects for rigorous value-added assessment over time.

We would also suggest that education philanthropies can make a unique contribution, in addition to that of states and school districts. In our survey of the research, we noted that the strongest findings to date arose from two experimental merit pay systems implemented in high schools in Israel (Lavy, 2002, 2004). Both of these systems were rank-order tournaments and involved substantial rewards for teachers. States and districts have been reluctant to implement

rank-order tournaments in part due to strong union opposition. By their very nature these are zero sum games and many assert that such incentive schemes discourage teacher collaboration and cooperation, to the detriment of overall school performance. Foundations, on the other hand, routinely award prizes to teachers, often in substantial sums. However, these awards are not implemented in a way that would permit evaluation of their incentive effects. Our survey of the literature suggests an opportunity for the philanthropic community both to do well and simultaneously add to a research knowledge base highly relevant to education policy.

References

- Aaronson, D., Barrow, L., & Sanders, W. (2003). Teachers and student achievement in Chicago public high schools. Chicago, IL: Federal Reserve Bank of Chicago.
- Altonji, J.G. & Pierret, C.R. (1996). Employer learning and the signaling value of education. Washington, DC: U.S. Department of Labor, Bureau of Labor Statistics.
- Angrist, J.D. & Guryan, J. (2005). Does teacher testing raise teacher quality? Evidence from state certification requirements. National Bureau for Economic Research Working Paper.
- Atkinson, A., Burgess, S., Croxon, B., Gregg, P., Propper, C., Slater, H., & Wilson, D. (2004) Evaluating the Impact of Performance-related Pay for Teachers in England. University of Bristol: Centre for Market and Public Organization.
- Armor, D., Conry-Oseguera, P., Fox, M., King, N. McDonnell, L., Pascal, A., Pauly, E. & Zellman, G. (1976). Analysis of the school preferred reading program in selected Los Angeles minority schools. Santa Monica, CA: RAND Corporation.
- Azordegan, J., Byrnett, P., Campbell, K., Greenman, J., & Coulter, T. 2005. Diversifying Teacher Compensation. Denver: Education Commission of the States
- Adnett, N. (2003). Reforming Teachers' Pay: Incentive payments, collegiate ethos and UK policy. *Cambridge Journal of Economics*, 27(1), 145–157.
- Baker, G. (1992). Incentive contracts and performance measurement. *Journal of Political Economy*, 100, 598-614.
- Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics*, 29(1), 37–66.
- Ballou, D. & Podgursky, M. (1993). Teachers' Attitudes Toward Merit Pay: Examining Conventional Wisdom. *Industrial and Labor Relations Review*, 47(1), 50-61.
- Ballou, D. & Podgursky, M. (2001). Let the market decide. *Education Next*, 1(1), 1-7.
- Ballou, D. & Podgursky, M. (1997). *Teacher Pay and Teacher Quality*. Kalamazoo, Michigan: W.E. Upjohn Institute for Employment Research.
- Ballou, D. (2001). Pay for performance in public and private schools. *Economics of Education Review*, 20(1), 51-61.
- Ballou, D. (2005). Value-added assessment: Lessons from Tennessee. In R. Lissitz (Ed.), *Longitudinal and Value Added Models of Student Performance*. Maple Grove, MN: JAM Press.
- Beer, M. & Cannon, M.D. (2004). Promise and peril in implementing pay-for-performance. *Human Resource Management*, 43(1), 3-48.

- Berhold, M. (1971). A theory of linear profit-sharing incentives. *Quarterly Journal of Economics*, 85(3), 460-482.
- Boyd, D., Grossman, P., Lankford, H., Loeb, S. (2006). How changes in entry requirements alter the teacher workforce and affect student achievement. *Education Finance and Policy*, 1(2), 176-216.
- Burgess, S., Crosson, B., Gregg, P., & Propper, C. (2001). The intricacies of the relationship between pay and performance of teachers: Do teachers respond to performance related pay schemes? Centre for Market and Public Organisation Working Paper, Series No. 01/35. U.K.: University of Bristol.
- Chamberlin, R., Wragg, T., Haynes, G., & Wragg, C. (2002). Performance-related pay and the teaching profession: A review of the literature. *Research Papers in Education*, 17(1), 31-49.
- Cohn, E. & Teel, S. (1992) Participation in a teacher incentive program and student achievement in reading and math. 1991 proceedings of the business and economic statistics section, American Statistical Association, Alexandria, VA.
- Cooper, S.T. & Cohn, E. (1997). Estimation of a frontier production function for the South Carolina educational process. *Economics of Education Review*, 16, 313-327.
- Clotfelter, C & H. Ladd (1996) Recognizing and Rewarding Success in Public Schools. In Ladd, H (Ed), *Holding Schools Accountable: Performance-related Reform in Education*, Washington D.C.: The Brookings Institution,
- Community Training and Assistance Center (2004). *Catalyst for Change: Pay for Performance in Denver Final Report*. Boston, MA: Author.
- Courty, P. & Marschke, G. (2003). Dynamics of Performance-Measurement Systems. *Oxford Review of Economic Policy*, 19 (2) (Summer), 268-284.
- Cullen, J.B. & Reback, R. (2006). Tinkering toward accolades: School gaming under a performance accountability system. NBER Working Paper #12286. Cambridge, MA: National Bureau for Economic Research.
- Dee, T. & Keys, B.J. (2004). Does merit pay reward good teachers? Evidence from a randomized experiment. *Journal of Policy Analysis and Management*, 23(3), 471-488.
- Deere, D. & Strayer, W. (2001). Putting schools to the test: School accountability, incentives, and behavior. Texas A&M University, unpublished.
- Delisio, E.R. (2003). Pay for performance: What are the issues?, *Education World*.
- Dixit, A. (2002). Incentives and Organizations in the Public Sector. *Journal of Human Resources*, 37(4), 696-727.
- Eberts, R., Hollenbeck, K., & Stone, J. (2002). Teacher performance incentives and student outcomes. *Journal of Human Resources*, 37(4), 913-927.

- Ehrenberg, R., McGraw, M., & Mrdjenovic, J. (forthcoming). Why do field differentials in average faculty salaries vary across universities? *Economics of Education Review*.
- Figlio, D. (2003). Testing, crime and punishment. University of Florida, unpublished.
- Figlio, D. & Kenny, L. (2003). Do individual incentives boost student performance? Mimeograph. University of Florida.
- Figlio, D. & Getzler, L. (2002). Accountability, ability and disability: Gaming the system?. National Bureau for Economic Research Working Paper 9307. Cambridge: NBER.
- Figlio, D. & Winicki, J. (2005). Food for thought? The effects of school accountability plans on school nutrition. *Journal of Public Economics*, 89, 381-394.
- Glazerman, S., Silva, T., Addy, N., Avellar, S., Max, J., McKie, A., Natzke, B., Puma, M., Wolfe, P., Gresler, R.U. (2006). Options for studying teacher pay reform using natural experiments. Washington, DC: Mathematica Policy Research, Inc.
- Glewwe, P., Ilias, N., & Kremer, M. (2004). Teacher incentives. Mimeograph. Cambridge, MA: Harvard University.
- Goldin, C. (2003). The human capital century: Has U.S. leadership come to an end? *Education Next*, Winter, 73-78.
- Goldhaber, D. & Anthony, E. (2006). Can teacher quality be effectively assessed? National Board Certification as a signal of effective teaching. Washington, DC: Urban Institute.
- Goldhaber, D., Hyung J., DeArmond, M., & Player, D. 2005. Why Do So Few Public School Districts Use Merit Pay? University of Washington Working Paper.
- Goodnough, A. (1999). Answers allegedly supplied in effort to raise test scores. *New York Times*, December 8.
- Guthrie, J.W. (2006). Modern data systems for modern schools. University of Arkansas Working Paper.
- Guthrie, J.W. & Springer, M.G. (2006). Teacher pay for performance: Sound and lasting policy or another education reform fad. April 5th, p. 42.
- Guthrie, J.W., Springer, M.G., Rolle, A.R., Houck, E.A. (2007). *Modern Education Finance and Policy*. Mahwah, New Jersey: Allyn & Bacon.
- Haney, W. (2000). The myth of the Texas miracle in education. *Education Analysis Policy Archives*, 8(41).
- Hannaway, J. (1992). Higher Order Skills, Job Design, and Incentives: An Analysis and Proposal. *American Educational Research Journal*, 29(1), (Spring), pp. 3-21.

- Hanushek, E.A. (2003). The Failure of Input-Based Resource Policies. *Economic Journal*, 113(485), (February), F64-F68.
- Hanushek, E.A. & Rivkin, S. G. (2004). How to improve the supply of high quality teachers. In D. Ravitch (Ed.), *Brookings Papers on Education Policy: 2004* (pp.7-25). Washington, DC: Brookings Institution Press.
- Harvey-Beavis, O. (2003). Performance-related rewards for teachers: A literature review. Distributed at the Organization for Economic Cooperation and Development (OECD) Attracting, Developing and Retaining Effective Teachers Conference in Athens, Greece.
- Hassel, B.C. (2002). Better pay for better teaching: Making teacher pay payoff in the age of accountability. PPI Policy Report. Washington, DC: Progressive Policy Institute.
- Hatry, H., Greiner, J., & Ashford, B. (1994). *Issues and Case Studies in Teacher Incentive Plans*. Washington, DC: Urban Institute Press.
- Hein, K. (1996). Raises fail, but incentives save the day. *Incentive*, 170, 11.
- Heneman, R. & Ledford, G.E. (1998). Competency pay for professional and managers in business: A review and implications for teachers. *Journal of Personnel Evaluation in Education*, 12(2), 103-121.
- Holmstrom, B. & Milgrom, P. (1991). Multitask principal agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organizations*, 7, 24-52.
- Hoxby, C.M. (2002). Would school choice change the teaching profession? *Journal of Human Resources*, 37(4), 846-891.
- Hoxby, C.M. & Leigh, A. (2004). Pulled away or pushed out? Explaining the decline of teacher aptitude in the United States. *American Economic Review*, 93(2), 236-240.
- Jacob, B. (2005). Testing, accountability, and incentives: The impact of high-stakes testing in Chicago Public Schools. *Journal of Public Economics*, 89(5-6).
- Jacob, B. & Lefgren, L. (2005). Principals as Agents: Subjective Performance Measurement in Education. National Bureau for Economic Research Working Paper 11463. Cambridge: NBER.
- Jacob, B. & Levitt, S. (2005). Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *Quarterly Journal of Economics*, 118 (3).
- Jensen, M. & Meckling, W. (1976). Theory of the firm: Managerial behavior, agency costs, and ownership structure. *Journal of Financial Economics*, 3, 305-360.

- Kahlenberg, R.D. (1990). The history of collective bargaining among teachers. In J. Hannaway and A. Rotherham (Eds.), *Collective Bargaining in Education: Negotiating Change in Today's Schools*. Cambridge, MA: Harvard University Press.
- Kandel, E. & Lazear, E. (1992). Peer pressure and partnerships. *Journal of Political Economy*, 100, 801-817.
- Kane, T.J., Rockoff, J.E., & Staiger, D.O. (2005). Identifying effective teachers in New York City. Paper presented at NBER Summer Institute.
- Kelley, C. (1999). The motivational impact of school-based performance awards. *Journal of Personnel Evaluation in Education*, 12(4), 309-26.
- Kelley, C., Odden, A., Milanowski, A., & Heneman, H. (2000). The motivational effects of school-based performance awards. Consortium for Policy Research in Education Policy Brief. Philadelphia, PA.
- Kelley, C., Heneman, H., & Milanowski, A. (2000). School-based performance award programs, teacher motivation and school performance: Findings from a study of three programs. Consortium for Policy Research in Education Report Series RR-44. Philadelphia, PA.
- Kelley, C., Heneman, H., & Milanowski, A. (2002). Teacher motivation and school-based performance rewards. *Education Administration Quarterly*, 38 (3), 372-401.
- Kerchner, C.T., Koppich, J.E., & Weeres, J.G. (2003). *United Mind Workers: Unions and Teaching in the Knowledge Society*. San Francisco, California: Jossey-Bass, Inc.
- Kershaw, J.A. & McKean, R.N. (1962). Teacher shortages and salary schedules. A RAND Research Memorandum. Santa Monica, CA: RAND Corporation.
- Koedel, C. & Betts, J. (2005). Re-examining the role of teacher quality in the education production function. University of California – San Diego Working Paper.
- Koretz, D., et al. (1999). Perceived Effects of the Kentucky Instructional Results Information System (KIRIS). Santa Monica, CA: RAND Corporation.
- Koretz, D. (2002). Limitations in the use of achievement tests as measures of educators' productivity. *Journal of Human Resources*, 37(4), 752-777.
- Koretz, D. & Barron, S.I. (1998). *The Validity of Gains in Scores on the Kentucky Instructional Results Information System (KIRIS)*. Washington, DC: Educational Resources Information System.
- Ladd, H. (1999). The Dallas school accountability and incentive program: an evaluation of its impacts on student outcomes. *Economics of Education Review*, 18(1), 1-16.
- Lavy, V. (2002). Evaluating the effect of teachers' group performance incentives on pupil achievement. *Journal of Political Economy*, 110(6), 1286-1317.

- Lavy, V. (2004). Performance Pay and Teachers' Effort, Productivity and Grading Ethics. National Bureau for Economic Research Working Paper 10622. Cambridge: NBER.
- Lazear, E. (2003). Teacher Incentives. *Swedish Economic Policy Review*, 10, 197-213.
- Lazear, E. (2002). Performance pay and productivity. *American Economic Review*, 90, 1346-1361.
- Lockwood, J.R., Doran, H & McCaffrey, D.F. (2003). Using R for estimating longitudinal student achievement models. *The R Newsletter*, 3(3), 17–23.
- Lockwood, J.R., McCaffrey, D. F., & Louis, T.A. (2002). Uncertainty in rank estimation: Implications for value-added modeling accountability systems. *Journal of Educational and Behavioral Statistics*, 27(3), 255–270.
- McCaffrey, D.F., Lockwood, J.R., Koretz, D., Louis, T.A. & Hamilton, L.S. (2004). Models for value-added modeling of teacher effects. *Journal of Education and Behavioral Statistics*, 29(1), 67-101.
- McCaffrey, D.F., Lockwood, J.R., Koretz, D.M., & Hamilton, L.S. (2003). Evaluating value added models for teacher accountability, MG-158-EDU. Santa Monica, CA: RAND Corporation.
- Malanga, S. (2001). Why merit pay will improve teaching. *City Journal*, 11(3).
- Marsden, D. & Belfield, R. (2006). Pay for performance where output is hard to measure: The case of performance pay for school teachers. In B.E. Kauffman and D. Lewin (Eds.), *Advances in Industrial and Labor Relations*. London, UK: JAI Press.
- Milken Family Foundation (2005). *Understanding the teacher advancement program*. Santa Monica, CA: Milken Family Foundation.
- Moehlman, A.B. (1927). *Public School Finance*. New York: New York: Rand McNally & Company.
- Murnane, R.J. (1975). *The impact of school resources on the learning of inner city children*. Cambridge, MA: Ballinger Press.
- Murnane, R.J. & Cohen, D. (1986). Merit pay and the evaluation problem: Why most merit pay plans fail and few survive. *Harvard Education Review*, 56(1), 1-17.
- Murnane, R. J., & Olsen, R.J. (1990). The effects of salaries and opportunity costs on length of stay in teaching: Evidence from North Carolina. *Journal of Human Resources*, 25(1), 106–124.
- National Commission on Teaching and America's Workforce (1996). *What Matters Most: Teaching for America's Future*. Washington, DC: Author.
- Odden, A. & Kelley, C. (1996). *Paying teachers for what they know and do: New and smarter compensation strategies to improve schools*. Thousand Oaks, CA: Corwin Press.

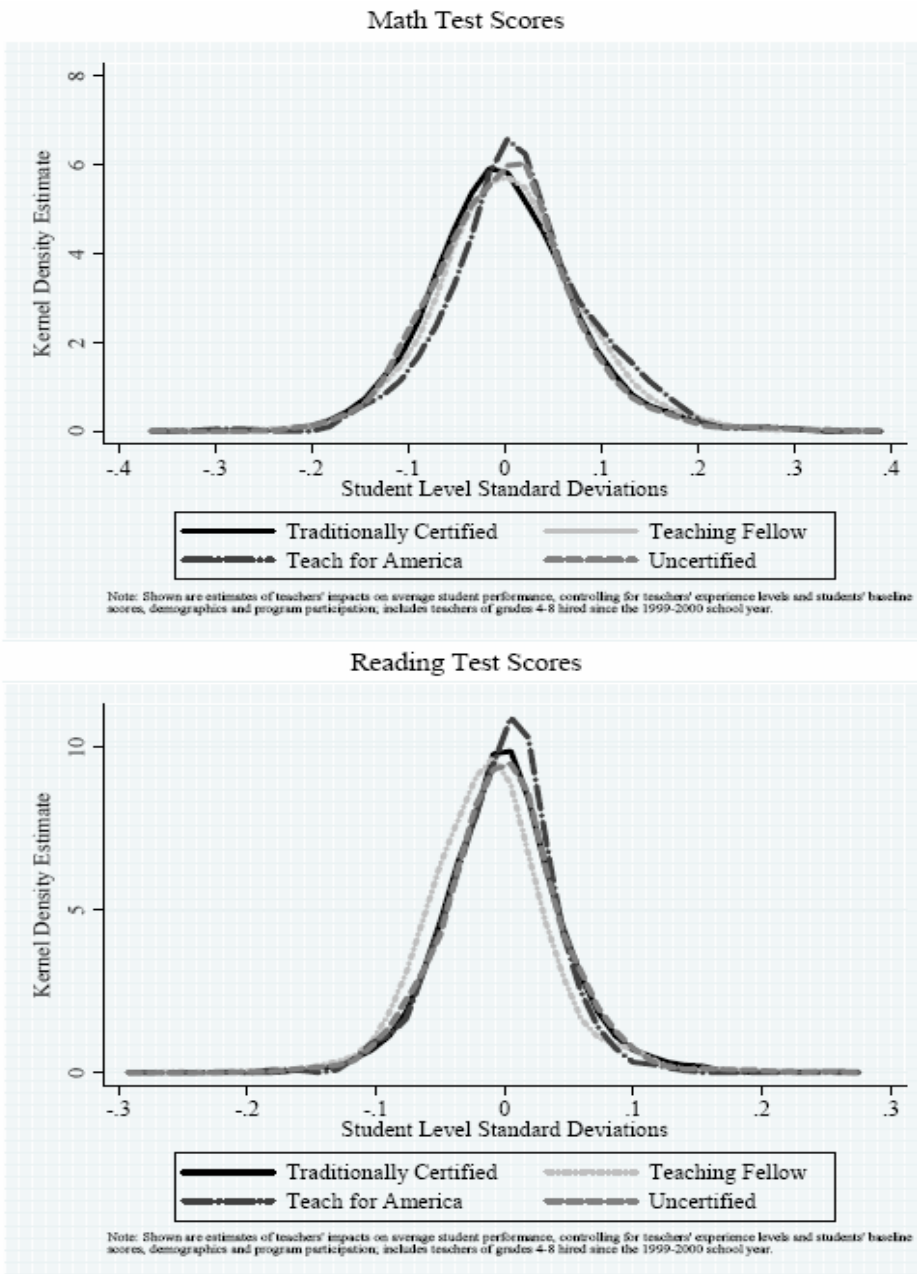
- Odden, A., Kelley, C., Heneman, H., & Milanowski, A. (2001). Enhancing teacher quality through knowledge and skills-based pay. Consortium for Policy Research in Education. Philadelphia, PA.
- Peabody, Z. & Markley, M. (2003). State May Lower HISD Rating; Almost 3,000 Dropouts Miscounted, Report Says. Houston Chronicle, June 14, A1.
- Podgursky, M. (2006). Teams versus Bureaucracies: Personnel Policy, Wage-Setting, and Teacher Quality in Traditional Public, Charter, and Private Schools. Education and Policy Analysis Archives.
- Podgursky, M., Monroe, R., & Watson, D. (2004). The academic quality of public school teachers: An analysis of entry and exit behavior. *Economics of Education Review*, 23(5), 507-518.
- Prendergast, C. (1999). The provision of incentives in firms. *Journal of Economic Literature*, 37, 7-63.
- Protsik, J. (1995). History of teacher pay and incentive reform. Washington: Educational Resources Information Center.
- Rivkin, S., Hanushek, E.A., & Kain, J.F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2), 417-458.
- Ross, S. (1973). The economic theory of agency: The principal's problem. *American Economic Review*, 63(2), 134-139.
- Sander, W.J., Ashton, J.J., & Wright, S.P. (2005). Comparison of the effects of NBPTS-certified teacher with other teachers on the rate of student achievement progress. Cary, NC: SAS Institute Inc.
- Sanders, W.J. & Horn, S.P. (1994). The Tennessee Value-Added Assessment System: Mixed-model methodology in educational assessment. *Journal of Personnel Evaluation in Education*, 8(3): 299–311.
- Sclafani, S. & Tucker, M.S. (2006). Teacher and principal compensation: An international review. Washington, D.C: Center for American Progress.
- Sharpes, D.K. (1987). Incentive pay and the promotion of teaching proficiencies. *The Clearinghouse*, 60, 407-410.
- Smith, S. & Mickelson, R. (2000). All that glitters is not gold: School reform in Charlotte Mecklenburg. *Education Evaluation and Policy Analysis*, 22(2), 101-127.
- Tyack, D.B. (1975). *The One Best System: A History of American Urban Education*. Cambridge, MA: Harvard University Press.
- Vyrastekova, J., Onderstal, S., & Konig, P. (2006). Team incentives in public organisations: An experimental study. CPB Discussion Paper No. 60. The Hague: Netherlands Bureau for Economic Policy Analysis.

Wong, K.K., Guthrie, J.W., & Harris, D.N. (2004). *A Nation at Risk: A Twenty-Year Reappraisal*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Wright, S.P., Horn, S.P., & Sanders, W.L. (1997). Teacher and classroom context effects on student achievement. *Journal of Personnel Evaluation in Education*, 11, 57-67.

Figure 1

Variation in Teacher Effectiveness by type of Teacher Certificate:
New York City Public Schools, 1998-99 – 2004-05



Source: Kane, Rockoff, Staiger (2006, Fig. 6)

Table 1

Major Teacher Performance Pay Programs

<u>Name of Plan</u>	<u>Target</u>	<u>Size of Bonus</u>	<u>Size of Program</u>	<u>Year of Inception</u>	<u>Funding Source</u>
Denver's Professional Compensation System for Teachers (ProComp)	Teacher Award Program	<u>Knowledge and Skills</u> : \$1000 tuition credit for Professional Development Coursework; 2% salary index bonus for completing courses and demonstrating skills (\$659); 9% salary index bonus for NBPST certification (\$2,967).	Pilot program operated in 16 schools	Pilot program operated from 1999 through 2004	Scaled-up program locally-funded following a 1 mill levy approved by taxpayers
		<u>Professional Evaluation</u> : Salary increase of 3% index for satisfactory evaluation of non-probationary teacher (\$989).	Scaled-up program implemented district-wide	Scaled-up program implemented district-wide in 2005	
		<u>Student Growth</u> : 3% sustainable increase for CSAP goal completion (\$989); bonus of 2% index for “distinguished” schools (\$659); bonue of 1% for meeting one of two goals (\$330).			
		<u>Market Incentives</u> : 3% index bonus for Hard to Staff (\$989) or Hard to Serve (\$989) assignments			
		<u>Total Bonus Range</u> : \$330-7582			

Texas' Governor's Educator Excellence Award Grants	Includes Three Scho	School-based awards range from \$40,000 to \$290,000 per year based on student enrollment	The three programs will include approximately 1300 schools and \$330 million per annum	Program was announced in 2006	A combination of state and federal funds
		Texas Education Agency recommends Individual teacher awards range from \$3,000 to \$10,000			

<i>Texas Educator Excellence Grant</i>	School Award Program	School-based awards range from \$40,000 to \$290,000 per year based on student enrollment	1,163 school are eligible during the 2006-2007 school year	Program was implemented during the 2006-2007 school year	Program is funded through state appropriations
	Schools must be in top half of schools in percentages of economically disadvantaged students.	75% of award must be paid to full-time classroom teachers based on a variety of objective measures of student performance (Part I)	\$100 million annually through 2009		
		25% to all school personnel, including principals, and/or professional development activities (Part II funds)			

	School Award Program	School-based awards range from \$60,000 to \$220,000 per year based on student enrollment.	Approximately 100 schools are eligible	Pilot program was implemented in 2006	Pilot program is funded through federal appropriations
<i>Governor's Educator Excellence Awards Pilot</i>	Schools must be in top third of schools in percentages of economically disadvantages students and have performance rating of either Exemplary or Recognized, or must in the top quartile of TEA's Comparable Improvement measure.	75% of award must be paid to full-time classroom teachers based on a variety of objective measures of student performance (Part I)	\$10 million annually through 2008.		
		25% to all school personnel, including principals, and/or professional development activities (Part II funds)			

	District Award Program	District-based award that is contingent upon district and school size			
<i>District-Level Grants Program (To Be Named)</i>	All school districts are eligible	60% of funds to directly award classroom teachers 40% of funds go to other personnel stipends and/or TAP implementation	\$230 million annually through 2010	Program will be implemented in 2008 school year	State funded
Florida's Special Teachers Are Rewarded (STAR)	Teacher Award Program	At least 5 percent of a teacher's base salary	\$147.5 million	Approved in 2006-07 budget	State funded by the Florida Education Finance Program (FEFP)
	At least 25% of a districts instructional staff will receive awards				
	Remaining funds to award excellence among the faculty				

Minnesota's Q-Comp	Schools receive funds to award teachers for excellence in student achievement.	Districts receive \$260 per student to implement program	\$86 million	State funded	
			Currently in 22 districts with 134 additional districts expect by 2008 school year		
Milken Family Foundation's Teacher Advancement Program (TAP)	Individual Teachers	<u>Master Teachers</u> : \$5,000 to \$11,000	9 states totaling approximately 50 school districts	1999	Private Family Phianthropic Foundation
		<u>Mentor Teachers</u> : \$2,000 to \$5,000	10 additional states are actively pursuing implementation		
		There is a range in the size of erformance bonuses - TAP recommends school average bonus of \$2,500 per teacher			

Table 2
Quantitative Studies of the Causal Effect of Teacher Incentive Programs on Measures of Student Achievement

Study	Sample	Time Span of Study	Type of Teacher Incentive	Size of Incentive (per teacher)	Outcome Variable	Results
Ladd (1999) Clotfelter and Ladd (1996)	Dallas grade 7 schools relative to other Texas urban districts ²⁰	1991-1995	School-wide (tournament)	\$1000	Math and reading test scores, dropout rates	Positive
Eberts, et.al. (2002)	2 MI alternative high schools (1 treatment, 1 control)	1994/95 – 1998/99	Individual	Up to 20 percent of base pay	Course completion rates, pass rates, daily attendance, GPA	Mixed
Lavy (2002)	Israel, high schools	1993-95 – 1996-97	School-wide (tournament)	\$200-\$715	test scores, pass rates, dropout rates, course-taking	Positive
Lavy (2004)	Israel, high schools	1999-2001	individual (tournament)	\$1750 - \$7500+ ²¹	pass rates and test scores	Positive
Glewwe, et.al. (2004)	Primary schools, rural Kenya	1997-1999	School-wide	Up to 43 percent of monthly salary	Grade 4, 8 test scores	Mixed
Atkinson, et.al. (2004)	UK High Schools	1997-2002	Individual	> 9 percent in salary base	English, Science, Math assessments	Positive
Figlio and Kenney (2006)	NELS-88 matched, to FK survey or 1993-94 SASS, 12 th grade Public and private	1993	individual	Varied within sample	12th grade, composite reading, math science and history score	Positive

20. Incentive applied to all schools but data limitations only permitted examination of grade 7 effects.

21. These are winnings per class. However, a teacher could enter multiple classes.

Table 3

Teacher Salary Schedules and Teacher Incentive Pay in Traditional Public, Charter, and Private Schools

(standard errors in parentheses)

	<u>Traditional Public</u> (%)	<u>Charter</u> (%)	<u>Private</u> (%)	<u>Non-religious Regular School</u> (%)
Is there a salary schedule for teachers in this school?	96.3 (0.29)	62.2 (0.72)	65.9 (1.24)	45.1 (5.60)
Does this school currently use pay incentives such as cash bonuses, salary increases, or different steps on the salary schedule to reward:				
NBPTS Certification?	8.3 (0.37)	11.0 (0.43)	9.6 (0.88)	14.8 (5.5)
Excellence in Teaching?	5.5 (0.35)	35.7 (0.65)	21.5 (0.93)	42.9 (5.5)
Completion of in-service professional development?	26.4 (0.70)	20.5 (0.56)	18.7 (0.88)	26.0 (5.67)
Recruit or retain teachers in fields of shortage?	10.4 (0.464)	14.9 (0.54)	7.9 (0.61)	15.0 (3.40)

Source: 1999-00 Schools and Staffing Surveys, reported in Podgursky (2006)