Why Do So Few Public School Districts Use Merit Pay?

Dan Goldhaber
Michael DeArmond
Daniel Player
Hyung-Jai Choi

information contact
Dan Goldhaber
206-685-2214
dgoldhab@u.washington.edu

June 28, 2005

Center on Reinventing Public Education
Daniel J Evans School of Public Affairs• University of Washington, Box 353055 • Seattle WA 98195-3055
206.685.2214 • www.crpe.org

CRPE working paper # 2005_5

A version of this paper was published in 2008, in *Journal of Education Finance*, 33(3): 262-289.

June 2005

Why Do So Few Public School Districts Use Merit Pay?

Abstract

The general theory of incentives suggests that public school teachers will face weak performance incentives because of the nature of their work and workplace: teaching is complex and multidimensional work; schools have multiple goals that are vague and poorly observed; multiple (and mobilized) stakeholders are interested in public education. All of this is a recipe for low incentives. Others suggest that incentives in education are low not because of the nature of teaching, but because of the high degree of unionization among public school teachers. Despite the ostensibly slim odds suggested by both views, a few school districts and states nevertheless offer performance incentives for teachers in the form of merit pay. In this paper we present a principal-agent model in the context of public schools to help explain the factors that affect district decisions about merit pay. The model includes the possibility that both the nature of teaching and the political costs of reform associated with unionization play a role in these decisions. It predicts that districts will be more likely to offer merit pay when they have more information about teacher performance (in effect altering the nature of teaching) and less likely to do so when their teachers are unionized. The model also suggests that, all else equal, teacher salaries will be higher in districts that offer merit pay. To test the model's predictions we analyze data from the U.S. Department of Education's Schools and Staffing Survey for 1999-00 and Census 2000 School District Demographics. Our results suggest that the political costs of reform affect district merit pay decisions and that teachers in merit pay districts earn more than their counterparts in non-merit pay districts. We find little evidence, however, in support of the hypothesis that more information about teacher performance makes merit pay more likely.

INTRODUCTION

In November 2005, citizens in Denver, Colorado will vote on whether or not to fund a new teacher pay system called the Professional Compensation System for Teachers (ProComp). On balance, if voters agree to the $25 million tax hike earmarked for the program, ProComp would put an end to the district's long-standing practice of paying teachers based solely on experience and education according to steps on a "single salary schedule." In broad strokes, the new plan rewards teachers for satisfactory evaluations from principals, additional training and education, and performance as measured by student achievement gains. The new system would also give extra pay to teachers who work in high-poverty schools or in hard-to-staff subjects like math and science.[1]

Although ProComp has multiple components, its pay-for-performance provision seems to have taken on master status – indeed, ProComp is commonly referred to as Denver's "merit pay" plan. This tendency to focus on ProComp's pay-for-performance provisions should come as no surprise. Attempts to link teacher pay to performance have a long and controversial history in education. On the one hand, economic theory suggests that merit pay could be a successful way to improve schools by attracting more able people to teaching and motivating them to be more productive. On the other hand, many educators (and researchers) argue that tying compensation to performance cannot work in schools because, among other things, teachers engage in multiple goals (which one should be rewarded?) and, moreover, because successful teaching is notoriously difficult to quantify.

Research on merit pay is equally mixed. While efforts by school districts to offer teachers merit pay have been largely deemed unsuccessful in practice (Murnane and Cohen, 1986; Hatry, Geiner, and Ashford, 1994; Ballou, 2001; Goldhaber, 2002), other research lends support to the notion that merit pay may have some potential in education (Clotfelter and Ladd, 1996; Lavy

2002; Dee and Keys 2004, Figlio and Kenny, 2005). But even if one thinks the jury is still out on merit pay's worthiness, it is certainly fair to say that it has a decidedly poor reputation in education (e.g., see Farkas, Johnson, and Duffett. 2003).

There are two main schools of thought about merit pay's unhappy history in education. First, as already suggested, there is the idea that merit pay is simply ill suited to the work that teachers do. In an influential article, Murnane and Cohen (1986) suggest that merit pay and education do not mix because the complex work that teachers do is difficult to evaluate. Without clear measures and criteria for judging success, decisions about rewarding performance are, at best, subjective and, at worst, unworkable. They also suggest that merit pay is problematic because it raises the potential for dysfunctional (or, as they call it, opportunistic) behavior: that is, teachers may end up focusing only on tasks that are rewarded by a merit pay plan at the expense of additional goals or tasks valued by the public (e.g., promoting citizenship, or reducing drug use or violence). Furthermore, at its worst, merit pay may have a demoralizing and counterproductive effect on the work place, corroding teacher collegiality by introducing competition. In sum, this is the view that there is something about the nature of teaching and schooling that makes the effective use of merit pay in public education unlikely.

By contrast, Ballou (2001) argues that there is nothing inherent in teaching and schooling that makes merit pay a poor fit. As evidence he points to the prevalence of merit pay in the nation's non-sectarian private schools. In 1993, for example, he finds that while only 12.3 percent of public school districts reported using merit pay, the number for non-sectarian private schools was almost three times higher—35.4 percent. Ballou also finds that private schools provide relatively large merit rewards—worth almost 10 percent of earnings, on average—and that they do so with discretion (that is, it appears that not all private school teachers get rewards). The fact that private schools use merit pay more often than public schools suggests that its failure

in public education may not be related to the nature of teaching and schooling per se. After examining the incidence of merit pay, the survival of merit pay plans, and the size of merit awards in districts where there are different degrees of union influence (i.e., those with collective bargaining agreements versus those with meet and confer agreements versus those with no labor agreement), Ballou concludes that the real problem with merit pay in education may be teachers unions: on all three measures, merit pay plans fare poorly where teachers unions have more influence. For Ballou, the political costs created by teachers unions have more to do with merit pay's problems in public education than any inherent aspect of teaching.

Of course, neither of these viewpoints – let's call them the "nature of teaching" hypothesis and the "political cost" hypothesis – is quite so straightforward. For example, although teaching performance is indeed difficult to measure, researchers and policymakers have recently explored new ways to separate out the effect teachers have on student achievement. Although methodologically challenging, these complex statistical techniques, along with longitudinal student test scores, may eventually help policymakers identify teacher value-added achievement effects (See McCaffrey et al., 2004).

Likewise, the union story is subtler than it seems. As Denver's experience suggests, union opinion about merit pay is not monolithic. When local affiliates explore alternative compensation systems, the nation's largest teachers union, the National Education Association (NEA), appears to subscribe to a "don't ask, don't tell" policy, even though it rejected a resolution in 2000 that would have endorsed experiments in performance-based pay. Sandra Feldman, the former president of the nation's second largest teachers union, the American Federation of Teachers (AFT), has written that pay reforms could include both salary increases and rewards for "different roles, responsibilities, skills, and yes, results" (Feldman, 2004).

Moreover, as Ballou notes, different local unions have different amounts of power depending on the kinds of labor agreements they have with school districts.

Even with these caveats in mind, Murnane and Cohen's "nature of teaching" hypothesis and Ballou's "political cost" hypotheses offer logical, competing explanations for why so few public school districts use merit pay. Ballou's political argument is intriguing and fits with existing theory on union wage determination, which suggests that unionized wages fit the preferences of the average worker, not the marginal worker with extra skills or quality (e.g., Freeman and Medoff, 1984). The general theory of incentives is on Murnane and Cohen's side. Dixit (2002), for example, argues that it is exactly the job characteristics that Murnane and Cohen pay attention to (the complex and multidimensional tasks of teaching; the multiple goals that are vague and poorly observed), coupled with the presence of multiple and mobilized stakeholders, that create low incentives in the public sector. Given the nature of teaching and public education, Dixit concludes, "We should expect to see weak explicit incentives, many constraints, and evaluation by evidence that the rules were followed" (Dixit, 2002, p. 721). In a recent survey of the theory of incentives, Prendergast (1999) makes a complementary counter-point: piece-rate merit rewards (i.e., rewards for units of output) are most likely to work in "simple" jobs where performance measures are easily available. Either way, the theoretical deck appears stacked against merit pay in public education.

Given that we should generally expect weak performance incentives in education, one might argue that recurring calls for linking teachers' pay to performance do not hold much promise, even in the face of long-standing criticisms about the single salary schedule (e.g., that the salary schedule provides few incentives for performance and rewards teacher characteristics that are thought to be poor proxies for quality [Goldhaber, 2002; Goldhaber & Brewer, 1997; Hanushek, 1986, 1997]). And yet, numerous school districts still experiment with merit pay

(Hatry et al., 1994).[2] Today, reformers in Denver are charging ahead, and others in Minnesota, Idaho, Mississippi, and elsewhere may follow suit (Hoff, 2005). Interested outside observers of education continue to advocate merit pay as an important school reform (e.g., Committee for Economic Development, 2004; The Business Roundtable and National Alliance of Business, 2000). All of these examples run counter to what we might expect. Why, given the odds, would a district abandon the familiar ground of the single salary schedule and adopt a pay-for-performance plan? In spite of all the notice paid to the issue, there is practically no empirical evidence that addresses this question.

In this paper we add to the merit pay discussion by presenting a simple principal-agent model in the context of public schools to explain district decisions about offering merit pay.[3] Our aims are two-fold. First, we hope to contribute to the literature on teacher compensation by grounding our work – both our hypothesis and variable specification – in a formal theoretical model that incorporates questions about the nature of teaching as well as the political costs associated with the reform.[4] Second, and more practically, we hope to shed some light on key contextual factors, in both policy and politics, which may influence district pay decisions.

Our model predicts that districts will be more likely to use some form of pay for performance in a policy environment where they have more information about teacher performance (in effect altering the nature of teaching), and less likely where the political costs of reform are high. It also suggests that teacher salaries will be higher in districts that offer merit pay. On the whole, we find little support in our empirical analysis for the idea that more performance information increases the likelihood of merit pay. We find far more evidence supporting the notion that political costs play a key role in explaining the decision to use merit pay, and we find, consistent with our theoretical model, that merit pay may lead to higher average salaries. These results suggest that large-scale experiments with merit pay are unlikely

without fundamental changes in the political dynamic surrounding management-labor relations in public education.

The paper is arranged as follows: first, we present a principal-agent model in the terms of public education and intuit testable hypotheses about when would expect to see (or not see) merit pay. Second, we describe our methods and data. Third, we present results from our empirical tests. Fourth, we offer some concluding thoughts on policy implications and further research.

SCHOOL DISTRICT DECISIONS ABOUT PERFORMANCE INCENTIVES

As both the "nature of teaching" hypothesis and the "political cost" hypothesis suggest, we should expect performance incentives in public education to be relatively rare. On the whole, the evidence shows as much. In 2000, for example, only 5.5 percent of public school districts reported using pay incentives, such as cash bonuses, salary increases, or bumps in the salary schedule, to reward "excellence in teaching" (in the years since the U.S. Department of Education began surveying districts on the issue, the proportion of districts using merit pay has generally hovered slightly higher, around 10 percent [Ballou, 2001]). Prior to examining the factors that might help explain why some districts decide to put aside strict adherence to the single salary schedule to experiment with merit pay, it is worth clarifying more generally what a district might hope to accomplish with merit pay.[5]

First, a district might hope that offering merit pay would affect the composition of the teacher workforce (i.e., who chooses to teach and stay teaching in the district). Theoretically, the higher marginal incentives associated with merit pay would attract higher-ability and less risk-averse teachers. Second, a district might also hope that offering merit pay would affect how teachers behave. Theoretically, the higher marginal incentives associated with merit pay would cause teachers to alter their efforts in the classroom, either in level (more effort), in type (more effort aligned with the district's academic goals), or both. Although realistically both factors are

likely to play a motivating role in the merit pay decision, for tractability we focus on the behavioral motivation. Accordingly, we can approach the merit pay question as a principal-agent moral hazard problem.[6]

Principal-Agent Theory in the Context of Public Education

Principal-agent theory considers relationships in which one actor (the principal) wants another actor (the agent) to act on his or her behalf. The relationship represents a contracting problem in which the principal must pay for the agent's effort, which produces an outcome that affects the principal's payoff. A key distinction in principal-agent theory is that it assumes the relationship between the principal and the agent includes various information asymmetries. In the moral hazard case, the information asymmetry is that the agent knows more about the service in question and about how the work contract is being fulfilled than the principal does. The central problem for the principal is to structure an incentive scheme that will persuade the agent to act according to the principal's aims so that the principal's expected utility of the payoff is maximized (Dixit, 2002; Laffont, 2002). Though rarely used in education, principal-agent theory offers insight into why a school district would choose to offer merit pay or not.

Districts, like most employers, face the exact moral hazard problem described above: they have less information about the work teachers do than teachers have themselves. Accordingly, we can use a principal-agent framework to model district decisions about compensation and, in particular, the conditions under which they would decide to abandon the single salary schedule and replace it with merit pay in hopes of affecting teacher effort.

In applying principal-agent theory to education, we make several simplifying assumptions. First, we assume that there is a single principal and single agent: we regard the *school district* as the principal and *individual teachers* as the agent. While this pairing is only

one of many possible principal-agent relationships found in school districts – for instance, there is a long list of other stakeholders (parents, teachers unions, taxpayers, local businesses, etc.) that might be thought of as principals or agents (Dixit, 2002) – this simplification makes the problem more tractable.[7]

The district in the model represents the key decision-making body in the organization (i.e., its school board) and is considered to present a unified front. In reality, district leadership is often fractured and far more complex. School boards squabble and may divide into voting blocks; other stakeholders have interests in teachers' work; and, depending on a district's governance structure, decision-making authority may be dispersed among one or more parties (e.g., a mayor or state agency). Though these entanglements suggest that future work should model multiple principals, at this stage we adhere to the single principal assumption.

We regard the teacher as a representative member of a homogenous population of teachers (importantly, the agent does not represent the teachers union). We recognize, however, that teachers within districts are not homogenous; they differ in terms of their fields of expertise, experience, and motivations.

Second, we assume that the district is interested in only one teacher activity that we call "teaching" and in one outcome that we call "student achievement." In reality, teachers pursue multiple goals that districts care about (e.g., increasing students' academic achievement, fostering their emotional and physical growth, preparing them for citizenship, etc.). Nevertheless, given today's policy and political interest in improving achievement for *all students* it is not unreasonable to portray the district as principally interested in teaching activity that leads to student achievement outcomes. We make no assumptions about whether or not the reward teachers receive for their effort is public information, although one might argue that a public reward could affect teacher effort differently than an inconspicuous reward. We also assume that

the contract between the district and teacher is a single-shot event. In practice, a multi-year merit pay program would be far more dynamic. Year to year implementation would likely be affected, for example, by teachers' expectations about the rewards they receive. That is, teachers may grow to expect a yearly reward of a certain size under the program, regardless of their performance or other related factors, and react angrily when that expectation is not met. No one likes to be demoted (Murnane & Cohen, 1986).

Third, we model the district's decision under the assumption that it does not care about teacher utility. We also assume that teachers are risk averse and that they do not care about student achievement; teachers seek to maximize their own utility, which is a function of their consumption and effort level. In Appendix A we consider alternative scenarios that complicate the matter by altering the district's regard for teacher utility and the teacher's regard for student achievement. In terms of the implementation decision, the results of these alternative scenarios differ only in degree, not in substance, from what is presented here. The following approach is adapted from Dixit (1996).

A Model of School District Merit Pay Decisions

We begin by describing the outcome of interest to the district: student achievement. Equation (1) specifies that student achievement x is the sum of teacher effort t and other random effects ε. As in Baker (2002), the ε in equation (1) could include acts of nature as well as the uncontrollable actions of others, including, for instance, parents and community members. From the teacher's perspective, all factors outside of her control are random. This abstraction reflects the reality that student outcomes depend on a host of factors that go beyond the teacher's "teaching" efforts (Hanushek, 1979). By recognizing that many factors feed into successes and failures in the classroom, only some of which have to do with teachers, these random effects

reflect the kinds of measurement and evaluation problems noted by Murnane and Cohen (1986) and many critics of merit pay. We assume that ε is normally distributed with mean 0 and variance σ^2.

$$x = t + \varepsilon \tag{1}$$

Teachers seek to maximize their utility, which varies positively with their net benefit from their contract with the district:

$$Max \quad U(w) = -\exp(-rw) \tag{2}$$

The functional form in equation (2) reflects teachers' risk averseness, with r representing the coefficient of constant risk aversion. The net benefit teachers receive from the contract, represented by w, is their income minus the cost of their effort measured in dollars. As teachers exert more teaching effort, we assume that the cost of effort grows, and so we represent the dollar equivalent of the cost of exerting teaching effort t as $\frac{1}{2}t^2$.

The district's net benefit is:

$$NB^D = bx - z - C \tag{3}$$

The marginal value of the student achievement output is assumed to be positive for the district, so $b > 0$. Equation (3) shows that the district faces two kinds of costs if it chooses to implement a merit pay plan. First, it must compensate the teachers for their effort under the contract. This payment is represented by z. Second, the district will face additional costs, represented by C. The cost C is a function of the administrative costs, A, associated with the plan and, following Ballou's (2001) hypothesis, the political costs, P, associated with union opposition to merit pay, $C = f(A,P)$. The district's problem is to maximize its expected net benefit:

$$\underset{t}{Max} \quad E[NB^D] = bt - z - C \tag{4}$$

We address the merit pay implementation decision in two stages. As a benchmark, we begin with the first-best case in which the district can perfectly observe the teacher's effort. We then address the second-best case, in which the teacher's effort is not directly observed by the district (or, as we refer to it from this point forward, is "unobservable").

First-Best: the Implementation Decision when Teacher Effort is Perfectly Observable

If both the district and the teacher can directly observe the teacher's effort, the two parties can write a contract directly contingent on the teacher making a certain amount of teaching effort t in exchange for a payment z (in this case, because the contract is made directly on the teacher's effort, the final result is the same regardless of the various scenarios presented in Appendix A that alter how the district values the teacher's utility and how the teacher values student achievement).

The expected return for the district if it pursues the merit pay contract is simply the gain in student achievement that stem from the teacher's effort minus the payment to the teacher and implementation costs. The district will implement the plan only when its net benefit is greater than zero.

If the district implements the plan, the teacher's utility would be:

$$U = -\exp[-r(z - \frac{1}{2}t^2)] \qquad (5)$$

Here the teacher's net benefit is the payment z she receives from the district minus the cost of her effort $\frac{1}{2}t^2$.

Again, the district and teacher's regard for each other and the teacher's attitude about students do not affect the outcome in the first-best scenario. If the district holds no regard for the teacher's utility, its decisions must still be made in light of the teacher's participation constraint.

And so, normalizing the benefits and costs of the status quo to be zero, if the district opts to use merit pay, it will set $z = \frac{1}{2}t^2$ so that the teacher will participate in the plan and exert effort. Given this constraint, the district will choose t such that its net benefit in (3) is maximized. (If the district cared about the teacher's utility, it would maximize the joint surplus. Either way, the result is the same: the district chooses the teacher's effort level so that the district's expected net benefit $bt - \frac{1}{2}t^2 - C$ is maximized.) The first-order condition for the maximization is $b = t$. The maximized district net benefit becomes:

$$NB_1^{D*} = b^2 - C - \frac{1}{2}b^2 = \frac{1}{2}b^2 - C \qquad (6)$$

The result is straightforward: the district will choose implementation only when its benefit in (6) exceeds the administrative and political cost—that is, when the net benefit is greater than zero.

Second-Best: the Implementation Decision when Teacher Effort is Unobservable

Now we consider the implementation decision under the more realistic assumption that the district cannot perfectly observe the teacher's effort. In this case, the district and the teacher write a contract that is not directly contingent on the teacher's effort, but on some student achievement output x that is observable to both parties. For tractability, we restrict the contract to a standard linear reward scheme based on the student-achievement output level. The payment to the teacher for output x is $\alpha x + \beta$, where the coefficient α is the marginal reward the teacher receives for unit increases in student performance x, and the coefficient β is the base salary needed to meet the teacher's participation constraint.

The district will maximize its own benefit, with the teacher's maximization as the constraint. The teacher's expected utility from making effort *t* is the payment he receives from the district minus the cost of his effort, which can now be written as:

$$-\exp\{-r\alpha t + \frac{1}{2}r^2\sigma^2\alpha^2 - r\beta + \frac{1}{2}t^2\} \quad (7)$$

For simplicity, we can think of the teacher's expected utility in equation (7) as $-\exp(-ry)$, where y replaces w from equation (2). As shown below in equation (8), y can be thought of as the teacher's certainty-equivalent (CE) income, or the certain payment that gives her the same utility as the expected value of the utility she would get from the actual distribution of outcomes under the merit pay contract. A teacher will only accept the contract if the utility of the expected value of the plan is greater than or equal to the utility of his status quo certain income level. Moreover, because the teacher is risk averse, the expected value of the merit pay program must exceed her CE income. In other words, the district must offer the teacher a compensating differential for risk, and therefore, all else equal, the teacher's expected compensation will rise if her district adopts a merit pay plan.

$$y = \alpha t - \frac{1}{2}r\sigma^2\alpha^2 + \beta - \frac{1}{2}t^2 \quad (8)$$

Given that the district uses the incentive scheme $\alpha x + \beta$, the teacher will choose t to maximize her CE income. The first order condition for the maximization is $\alpha - t = 0$ or $t = \alpha$. Substituting for the teacher's effort, her CE income becomes

$$y = \frac{1}{2}\alpha^2 - \frac{1}{2}r\sigma^2\alpha^2 + \beta. \quad (9)$$

Again, the district maximizes its benefit with the teacher's maximization as the constraint

$$\underset{\alpha}{Max}\ (b-\alpha)t - \beta - C = (b-\alpha)\alpha - \beta - C \quad (10)$$

The first order condition is $\alpha = \frac{1}{2}b$. By plugging $\alpha = \frac{1}{2}b$ into the teacher's CE income and the district's net benefit, the district's surplus becomes:

$$NB_2^{D*} = \frac{1}{4}b^2 - C - \beta \tag{11}$$

and the teacher's surplus becomes:

$$S_2^{T*} = \frac{1}{8}b^2(1 - r\sigma^2) + \beta \geq 0 \tag{12}$$

The district will choose to implement the plan if equation (11) is greater than zero. The maximized joint surplus in this case is:

$$\tilde{S} = \frac{1}{8}b^2(3 - r\sigma^2) - C \tag{13}$$

Propositions

We consider three propositions that follow from the above model. First, it is clear from equation (11) that, as in the first-best, the district will only pursue the contract when the benefits exceed the costs C, including administrative and political costs. Given our interest in Ballou's (2001) hypothesis about political cost and merit pay we arrive at:

> *Proposition 1*: District will be less likely to offer merit pay where the political costs of implementation are high.

In our empirical analysis we consider political costs in terms of union influence, reasoning that the greater the union influence in a district, the greater the political costs of using merit pay. We also consider merit pay usage in charter schools, which, by virtue of being free from many of the laws and regulations that govern traditional public schools, likely face different political and labor dynamics than do traditional public school districts.

The second proposition is related to Murnane and Cohen's "nature of teaching" hypothesis and the evaluation problem. Recall that the teacher's risk in the model involves her risk aversion, r, as well as how well the district is able to capture her effort, signaled by the variance σ^2 of the error term in equation (1). Also recall that we consider the error term ε in equation (1) to include acts of nature as well as the uncontrollable actions of others inside or outside of the school. These factors are random from the teacher's perspective. In the second-best, equation (12) shows that the teacher's surplus is a function of the variance σ^2. As the variance increases, the district will have to offer a higher transfer payment β in order to satisfy the inequality. In short, the district's costs rise and fall with the variance and, accordingly, so too does the probability that it will utilize merit pay. If the district were somehow able to lower the variance σ^2 by using more performance information that filters out some of the uncontrollable actions of others in the school, the teacher's risk of being punished for factors beyond her control would be less and the cost to the district would be lower. [8] In effect, we argue that the availability of performance information may alter the nature of teaching:

> *Proposition 2*: Districts that have more performance information about teachers will be more likely to use merit pay.

We test the idea that performance information may alter the "nature of teaching" by looking at state accountability systems, reasoning that districts in states with stronger accountability and information systems are likely to have more performance information about teachers than those in states with weaker systems.

Finally, as already noted, the teacher's assumed risk aversion suggests that merit pay programs are likely to be more costly, on average, than the certainty of the single salary schedule:

Proposition 3: On average, teacher pay will be higher in districts that use merit pay compared to those that do not.

ANALYTIC APPROACH AND DATA

We approach our analyses of Propositions 1 and 2 in two ways. First, we conduct a probit analysis to consider how political costs and performance information affect the probability that a district offers *formal* incentives for merit pay. Second, we consider how political costs and performance information affect the amount of *informal* merit pay districts offer teachers. We conduct similar analyses of charter schools. Our approach for investigating Proposition 3 on the relative costs of merit pay is a straightforward regression involving teacher-level pay data and district characteristics, including the provision of formal merit pay plans. In this section we present our empirical models and describe the data we use in the analyses.

Formal Merit Pay

We begin with a probit analysis to look at how political costs and performance information affect the likelihood that a district offers a formal merit pay plan. Our basic model is a latent-variable approach of the form:

$$M_n^* = P_n\delta_1 + C_n\delta_2 + X_n\delta_3 + \varepsilon_n \qquad (14)$$

where,

$M_n = 1$ if $M_n^* \geq 0$
$M_n = 0$ otherwise

In the above specification M_n is a binary measure of whether district n has a formal merit pay plan or not, P_n is a vector of performance information, C_n is a vector of the costs of reform facing the district, and X_n is a vector of community controls.

Given the model's predictions, we would expect that the coefficients in $\delta_1 > 0$ and the coefficients in $\delta_2 < 0$. That is, we expect merit pay to be more likely where there is more performance information and less likely where political costs are higher.[9]

Informal Merit Pay

Because the political costs of formal merit pay programs may, as Ballou (2001) suggests, be prohibitively large, it is reasonable to expect that districts may engage in less-formal, "under-the-radar" merit pay programs. In other words, teachers may receive monetary or non-monetary rewards for performance even where districts do not officially offer merit pay. A district may, for example, reward a particularly effective teacher with a favorable extra-curricular assignment or a summer school position for which they receive additional pay; alternatively, a district may reward teachers by crediting them with additional experience on the salary schedule, in effect allowing them to leapfrog up the schedule's steps (see, for example, Jupp, 2005). The point here is that districts and school leaders may use one or more of these tactics to reward teachers and performance without having to go through the steps of designing a formal merit pay program.

Suggesting that informal merit pay may exist is one thing; examining it with actual evidence is another. After all, by definition, districts will not keep readily identifiable records of informal rewards. As a result, if we want to look at informal merit pay, we are left having to estimate its size and presence through statistical techniques.

We gauge the amount of informal merit pay in districts using a two-stage estimation. In the first stage, we estimate the following equation for individual teacher *i* in district *n*:

$$S_{in} = E_{in}\lambda_1 + X_{in}\lambda_2 + ST_n\lambda_3 + \mu_i \qquad (15)$$

where S is the teacher's total teaching salary, E is a vector of dummies for the teacher's education level (MA, PhD, etc.), X is a vector of the teacher's experience and experience

squared, and ST is a vector of state dummies to account for the fact that state salary schedules differ from one another.[10] The logic behind equation (15) is that informal merit pay would be income that is not be explained by the factors that usually determine compensation under the single salary schedule (education and experience). In order to measure informal merit pay as deviations from the salary schedule, we predict the error $\hat{\mu}$ for each teacher. The relevant measure of informal merit pay for the second stage is the absolute value of the predicted error $\hat{\mu}$.

For the second stage we average the absolute values of the teacher predicted errors for each district. We then use these averages in a district regression of the form:

$$\overline{\hat{\mu}}_n = P_n\delta_1 + C_n\delta_2 + X_n\delta_3 + \varepsilon_n \qquad (16)$$

Except for the dependent variable, $\overline{\hat{\mu}}_n$, which measures deviation from the single salary schedule, equation (16) is similar to equation (14) with vectors of performance information (P_n) the costs of reform (C_n), and community characteristics (X_n). We estimate this second stage using OLS. We expect the signs on the coefficients to be the same as in the probit model ($\delta_1 > 0$ and $\delta_2 < 0$), but their interpretation is different. With the OLS analysis, the coefficients no longer represent changes in probability. Instead they reflect changes in the average size of the deviation from the single salary schedule, our measure of informal merit pay.

Clearly, some caveats are in order. First, it is important to note that the deviations measured by equation (15) could easily represent payments for any number of extra-salary schedule characteristics, ranging from the literally meritorious to the questionable (e.g., having a certain *alma mater* or family connection). Similarly, these deviations might simply represent the fact that some teachers work in districts with high costs relative to other districts in the same state. This is because the use of state dummies essentially treats teachers as if they are paid on a state salary schedule, something that, while generally true, will not hold in all cases. We attempt to control for the possibility that these deviations may simply be higher costs by including

urban/rural dummies and information on the median housing value and median income in our regressions. Even so, given the limitations of the data, the informal merit pay analysis should clearly be taken as no more than suggestive.

Data

Both our formal and informal merit pay measures come from the U.S. Department of Education's Schools and Staffing Survey (SASS). The SASS is a large sample survey of the nation's schools that has been administered in 1987-88, 1990-91, 1993-94, and 1999-00. We rely on data from the 1999-2000 School District and Teacher Questionnaires.[11] That year, the SASS sample included 5,465 public school districts and 56,354 public school teachers. The weighted response rates to the survey were high: 88.6 percent for school districts and 83.1 percent for teachers.[12]

Finding a formal merit pay measure in the SASS is fairly straightforward. District officials were asked if their district used any pay incentives such as cash bonuses, salary increases, or different steps on the salary schedule to reward "excellence in teaching." We use this question as a dummy variable in our probit analysis, indicating whether or not a school district offers merit pay (although the question avoided the term "merit pay"). This measure is less than ideal, however, as it does not allow us to measure the size of merit pay rewards. Moreover, it is likely that some districts may use "excellence in teaching" to refer extra work rather than actual classroom performance (Murnane & Cohen, 1986).

We use teacher-level data from the SASS on years of experience, degree level, and total teaching salary to arrive at the predicted errors in equation (15), which we then use in the second stage estimation in equation (16) to look at informal merit pay. The total teaching salary variable includes a teacher's self-reported base salary, earnings from extracurricular activities (like

coaching or supervising a student club), any merit pay or state supplements received, and any money earned from summertime teaching or other summertime school-based work.

Following Ballou (2001), we define political costs of merit pay reforms in terms of union influence: the greater the influence, the greater the political cost facing the district. The SASS includes information on whether or not districts have an agreement with a teachers union or organization. Districts are shown as having no agreement, a "meet and confer" agreement, or a collective bargaining agreement. We take these responses to represent a continuum of union influence, from weak (no agreement) to strong (a negotiated agreement), with the consultative "meet and confer" agreement representing something of a middle ground of influence. We also use data from the U.S. Department of Labor on state Right-to-Work (RTW) laws as another indication of union influence. In 2000, twenty-two states had such laws, which prevent labor-management agreements requiring workers to join a union as a condition of employment and, as a result, weaken union power. These RTW laws appear to have an important influence on unionization: in the 2000 SASS, only 33 percent of the districts in RTW states had collective bargaining agreements versus 81 percent in non-RTW states.

We control for other district and community characteristics that may be correlated with districts' pay decisions and thus bias our results by including data from the Census 2000 School District Demographics on the districts' median housing value, median income, adult education level (i.e., percent of adults with a bachelor's degree or higher), and size. In our OLS residual analysis of informal merit pay we also control for whether or not a district has a formal merit pay plan.

To get at the question of performance information (the availability of which, we argue, may alter the "nature of teaching"), we experiment with a variety of measures tied to state educational accountability systems. By the 1999-2000 school year, many states had or were

developing accountability systems that attempted to both define and measure statewide public education goals. We assume that the more well-defined these goals are, and the more sophisticated performance measures states use to measure progress toward them, the more districts are able to capture teacher contributions to student achievement and, so the logic goes, the more likely they are to use merit pay. Standards and accountability reforms, in other words, may alter the nature of teaching. Unfortunately, finding valid indicators of the degree of goal definition and measurement sophistication in state systems is difficult—little such information exists that is comparable across states. We explore several possibilities. First, we use an index from Carnoy and Loeb (2002) on the strength of state accountability systems. Carnoy and Loeb's Accountability Index (0-5) is based on state-survey data from the Consortium for Policy Research in Education (CPRE). The index gives states with no standards or tests the lowest score (0), and states with standards, tests (in primary and middle grades and a high school exit exam), plus sanctions and rewards for school-level performance the highest score (5) (Carnoy & Loeb, 2002).[13] From this point forward, we refer to the Carnoy and Loeb index as the Accountability Index.

Despite its utility, the Accountability Index poses some potential validity problems for our analysis because it is, in some ways, as much a measure of sanctions and rewards as it is of performance measurement. Fortunately, the same CPRE survey that forms the basis of the Accountability Index includes some limited information on how states measure progress. Based on these data we constructed our own 0-5 scale to capture the kinds of performance information used in each state. A high score on our index indicates that a state's performance information is relatively more sophisticated (and, we argue, likely alters the nature of teaching in that state). More particularly, our index gives those states with no system to measure progress (or a system under development) a zero; those states that measure school progress with an absolute

performance measure get a 1; those that measure progress with an absolute performance measure as well as growth from a baseline get a 2; those that measure progress with an absolute performance measure as well as a comparison to state averages get a 3; those that use absolute performance measures and comparisons to similar schools get a 4; those that use absolute performance measures and comparisons to predicted performance or value-added performance measures get a 5.

In this paper, we present results using three different measures of state accountability systems based on the above measures: first, we use the Accountability Index; second, we use a high-low binary version of the Accountability Index called High Accountability; and third, we use a 10-point index that combines the Accountability index with our index on performance information described in the preceding paragraph. We call this combined index the Performance Measure Adjusted Index. We also explored using two additional rating systems: Education Week's state-by-state ratings for Standards and Accountability from its Quality Counts 2000 report and state-by-state ratings from the Thomas B. Fordham Foundation's State of State Standards 2000 report. Neither of these alternatives had dramatically different results from the indices presented here (The coefficients from alternative specifications are available from the authors upon request).

One clear limitation of the indices we use is that they focus on information about school-level, not teacher-level, performance. Ideally, we would like information about teacher-level performance measures. Nevertheless, from a district's perspective, school-level measures are better than having no measures at all when it comes to shedding light on the performance of teachers within a school. And so, with all the appropriate caveats in mind, we reason that districts in states that score low on these three measures are more likely to confound the teacher's effect on achievement with other factors (namely student background), while those in higher

scoring states are more likely to have performance measurement systems that account, at least to some degree, for such factors.

To construct our sample we dropped districts in Washington D.C. because we do not have accountability measures for them. For districts missing control variables we impute values using mean value placement and include dummy variables identifying the imputed values in all regressions. Our final sample consists of 4689 districts for our formal merit pay analysis. Our OLS analysis of informal merit pay is restricted only to those districts that we were able to merge with teacher data, resulting in a smaller sample of 4070 districts.[14] Table 1 presents some descriptive statistics on all the districts in the formal merit pay sample. The teacher-level salary variables in Table 1 (base salary and total income from school), however, represent only districts that could be merged with teacher data. The districts and teachers included in Table 1 are weighted to be nationally representative.[15] The first column shows districts that offer formal merit pay; the second column shows districts that do not.

Table 1: Weighted Descriptive Statistics

Variable	Districts with Merit Pay (N=352) Mean	Districts without Merit Pay (N=4,337) Mean
Base Salary*	$36,453	$36,741
	(606.8)	(185.27)
Total Income From School*	$38,715	$38,495
	(507.3)	(193.4)
Accountability Index	2.58	2.36
	(0.10)	(0.01)
% High Accountability	44.44	44.49
	(2.60)	(0.31)
Performance Measure Adjusted Index	4.55	3.88
	(0.16)	(0.01)
% Right-To-Work	47.90	30.70
	(2.80)	(0.26)
% Collective Bargaining	49.63	65.00
	(3.24)	(0.80)
% Meet And Confer	8.23	5.60
	(2.34)	(0.41)
% Urban	8.31	5.95
	(0.80)	(0.01)

% Suburban	45.02	38.69
	(3.93)	(0.73)
% Rural	46.66	55.37
	(3.89)	(0.73)
% Minority	23.08	18.45
Students	(1.32)	(0.52)
% Subsidized-	38.92	36.18
Lunch Students	(1.91)	(0.50)

*Teacher-level salary variables

The sample statistics in Table 1 provide cursory evidence in support of the notion that political cost plays an important role in districts' merit pay choices. As the model predicts, there is an inverse relationship between merit pay and union influence: 48 percent of the districts with merit pay plans are located in RTW states, versus only 31 percent of those without merit pay plans.[16] Also, a smaller percentage of merit pay districts have collective bargaining agreements than districts without merit pay: 50 percent verses 65 percent.

The cross-tabulations in Table 2 show district merit pay usage by the number of districts with each type of labor agreement (no union, collective bargaining, and meet and confer). The numbers in the upper left of the cells in the table are for districts in RTW states; those in the lower left are for districts in Non-RTW states. As is apparent from the table, there are relatively few merit pay districts in our sample that have meet and confer agreements. Accordingly, the results involving meet and confer districts should be considered with a note of caution.

Table 2. Cross Tabulations of Merit Pay and Union Agreements Separated by Right To Work/Non-Right To Work States

	No Union	Collective Bargaining	Meet and Confer	**TOTAL**
No Merit Pay	861 / 361	541 / 2,079	211 / 121	**1,613 / 2,579**
Merit Pay	119 / 29	60 / 101	19 / 10	**198 / 140**
TOTAL	**980 / 390**	**601 / 2,198**	**230 / 131**	**1,811 / 2,719**

Returning to the sample statistics in Table 1, we find some support for the model's predictions about the relationship between performance information and merit pay. Two versions of the Accountability Index (the original Accountability Index, and the combined Performance Measure Adjusted Index measure) follow the model's predictions: merit pay districts appear more likely to be located in states with higher ratings on the indices, though not by much. Thus, it may be, as the model predicts, that the accountability and measurement regimes in these states alter the nature of teaching in a way that makes teaching more amenable to merit pay. We find a more mixed picture of the model's predictions about the relationship between average salaries and merit pay. The model predicts that salaries will be higher in districts with merit pay. While this is the case for the total salary variable, the means for base salary variable show the opposite: non-merit pay district teachers earn more money, though not by much.

Finally, the figures in Table 1 suggest that districts using merit pay are more likely to be in urban areas and to enroll poor and minority students. This could be due to a number of factors. The role played by the variance of the theoretical model's disturbance term implies that merit pay is more likely under circumstances where performance measures are less distorted or, seen another way, where teachers play a relatively larger role in determining student achievement. If, as appears likely based on empirical evidence (Coleman, 1990), teachers in districts enrolling large numbers of disadvantaged students play a relatively more important role in influencing achievement, then those same districts should, as Table 1 suggests, be more likely to use merit pay. Alternatively, merit pay might be more likely in these districts because, if we take poverty as a proxy for low achievement, the political costs of action may be lower in districts where achievement is lower. That is, politically controversial reforms might be more palatable in places where the status quo is clearly not working than in places where the situation does not appear as desperate.

Before taking a more careful look at the differences between districts that offer merit pay and those that do not, it is worth noting that the discussion so far has assumed that all of the independent variables in our regressions are exogenous. There are good reasons, however, to suspect that all of our accountability variables - the Accountability Index, the High Accountability measure, and the Performance Measure Adjusted Index - may not be exogenous. In particular, one could argue that the extent to which districts or states have strong accountability systems and the extent to which district are willing to offer merit pay may both be influenced by some third factor, e.g. how well students are performing in the state (i.e., low performance may drive both accountability *and* pay reforms). If this is the case, estimates of the impacts of the accountability variables will be biased. However, as we discuss in the next section, the results of a Durbin-Wu-Hausman test indicate that our estimates in both analyses do not suffer from endogeneity problems.

RESULTS

Probit Analysis of Formal Merit Pay

Table 3 reports the estimated marginal effects for the probit analysis of formal merit pay. Column one shows the results for the base model that uses the five-point Accountability Index, which gives states higher scores for having academic standards, for testing more grade levels, and for imposing sanctions and rewards for performance. Column two shows the results for the binary high-low version of this Accountability Index, called High Accountability. Column three shows the results for the Performance Measure Adjusted Index, which combines the Accountability Index with additional information on the kinds of performance information used in each state. All three specifications include school district demographic information and our

measures of political cost; that is, whether the district operates in a RTW state or under a collective bargaining or "meet and confer" agreement.

Prior to focusing on the performance and political cost variables, it is worth noting an interesting finding about school size. In all model specifications, the results clearly suggest that as average school size increases, districts are more likely to use merit pay. For every 100-student increase in the average school size, a district is one half of a percentage point more likely to use merit pay. From a theoretical perspective, it is unclear how organizational size would affect the merit pay decision. Kane and Steiger's (2002) finding that small schools can lead to distortions of test-based performance information suggests that merit pay might be less likely in smaller schools. District decision-makers who understand that small school produce noisy performance measures may choose to avoid merit pay. Likewise, the moral hazard problem may be less severe in small organizations, and so one might expect them not to bother with merit pay. Alternatively, however, one might argue that districts with smaller schools might actually have more (non-test) information about teacher performance and so should be more likely to use merit pay.

The results in Table 3 also suggest that enrolling more minority and low-income students does not significantly alter the probability of a district offering merit pay. The coefficient for percent minority students is in the expected direction; the coefficient for percent low-income students is in the opposite of the expected direction; and neither is statistically significant. (Again, the model indirectly predicts a greater likelihood of merit pay in districts with more disadvantaged students because schooling inputs are likely to be relatively more important than non-schooling factors in predicting achievement for these students).

Turning our attention to the variables designed to capture teacher performance information, we find little evidence that greater performance information increases the likelihood that districts use merit pay. The Accountability Index (column one) and the Performance

Measure Adjusted Index (column three) have the expected positive sign, indicating that districts may be more likely to offer merit pay when they face stronger state-level accountability. The sign on the High Accountability index is the opposite of what we expect. But none of the accountability measures are statistically significant at the 5 percent significance level.[17]

The above results, then, do not appear to support the "nature of teaching" hypothesis, at least as far as we have correctly measured districts' ability to capture teacher effort. However, as we discussed above, there is reason to be concerned about endogeneity when it comes to our accountability variables. Some third factor may affect both state accountability policy and the district-level merit pay decision. For example, both may be driven by a political preference in the state or by large numbers of low performing schools. If such a factor exists, our estimated equations will be biased.

To test for potential endogeniety, we used an instrumental variables approach with 2000 Census data on shifts in the power of political parties in state legislatures. We reasoned that a state might be more likely to adopt the legislation required for a strong accountability system and sophisticated performance measurement system if one political party dominated the state legislature. We also reasoned that if that party's domination were the result a marginal electoral gains – picking up one seat in the state house of representatives, for example - such consolidating electoral "flips" would be uncorrelated with individual school district decisions to offer unofficial merit pay. Our instruments jointly predicted the first stage; and, as already noted, Durbin-Wu-Hausman test results showed that endogeneity is not an issue. Moreover, when we re-estimated our equations using instrumental variables, the Heckman 2SLS results re-enforce our finding that a district's use of merit pay is not strongly related to the availability of teacher performance information, at least as far as our accountability variables adequately capture it.

There is far better evidence in support of the "political cost" hypothesis. In all three specifications, union power appears to act as a deterrent to merit pay use: districts without a collective bargaining agreements in non-RTW states are approximately 7.5 percentage points more likely to offer merit pay than districts in non-RTW states with collective bargaining. This represents a large change in the probability of having a merit pay system since only about 7.5 percent of the districts in our total sample have merit pay. Moreover, as the model predicts, the sign and size of the coefficients for "meet and confer" suggest that it too acts as a deterrent to merit pay, albeit a far weaker one than collective bargaining. We cannot explain the strange results on the interactions terms, which suggest that districts in RTW states with collective bargaining agreements are *more* likely to use merit pay than districts in RTW states without collective bargaining agreements.[18]

Table 3: Probability of Using Merit Pay
(Standard Errors in Parenthesis)

	Probit		
	(1)	(2)	(3)
Accountability Index	0.003		
	(0.003)		
High Accountability		-0.002	
		(0.008)	
Performance Measure Adjusted Index			0.002
			(0.001)
Right To Work	0.096	0.092	0.096
	(0.015)**	(0.014)**	(0.015)**
No bargaining agreement	0.076	0.077	0.075
	(0.019)**	(0.019)**	(0.019)**
Meet And Confer	0.049	0.046	0.049
	(0.031)	(0.031)	(0.031)
RTW*No Bargaining	-0.082	-0.075	-0.084
	(0.029)**	(0.028)**	(0.029)**
RTW*Meet And Confer	-0.078	-0.073	-0.078
	(0.036)**	(0.036)**	(0.036)**
District Enrollment > 1000	0.014	0.015	0.013
	(0.010)	(0.010)	(0.010)
Avg. School Enrollment/100	0.005	0.005	0.005
	(0.002)**	(0.002)**	(0.002)**
Urban District	0.017	0.018	0.016
	(0.014)	(0.015)	(0.014)
Suburban District	0.022	0.024	0.022
	(0.010)*	(0.010)*	(0.010)*
% Minority Enrollment	0.016	0.020	0.017
	(0.017)	(0.016)	(0.016)
% Eligible For Subsidized Lunch	-0.003	-0.002	-0.002
	(0.019)	(0.019)	(0.019)
Log Of Median House Value	0.024	0.025	0.024
	(0.013)	(0.013)	(0.013)
Log Of Median Income	-0.026	-0.025	-0.026
	(0.024)	(0.024)	(0.024)
% Of Pop. 25+ With At Least a BA	0.057	0.054	0.058
	(0.043)	(0.044)	(0.043)
Sample size	4689	4689	4689

* p < .05, ** p < .01

In order to get a further sense of the magnitude of the effect of political variables on the probability of a district offering merit pay, we use the coefficients from the base model (Table 3, column 1) to simulate the probabilities that districts offer merit pay given that they are located in RTW states or that they have collective bargaining agreements. The predicted probabilities in Table 4 show an important pattern. For districts located in states without RTW laws, the

presence of collective bargaining agreements has a large impact on whether the district offers merit pay. By contrast, in states with RTW laws, the difference in probabilities for districts with and without collective bargaining is much smaller; the political costs associated with a negotiated agreement appear far less in RTW states. Finally, the simulations also make it clear that districts in RTW states are more likely to have merit pay, whether or not they have collective bargaining.

Table 4. Simulated Probabilities of Having a Merit Pay Plan By Right to Work and Collective Bargaining Status

	Non-RTW	RTW
Collective Bargaining	.041	.137
No Collective Bargaining	.106	.120

Note: All means are statistically different from each other.

Analysis of Informal Merit Pay

Next we explore the model's propositions given the possibility that teachers are receiving additional compensation through informal merit pay systems. As with the probit analysis, the results of a Durbin-Wu-Hausman test indicate that our OLS model does not suffer from endogeneity problems.

We approach our informal merit pay analysis using a two-stage estimation (equations 15 and 16) and our sample is restricted to districts we were able to match with teacher data. The dependent variable represents the size of a district's deviation from the single salary schedule based on the total income from school of teachers in the district. Prior to focusing on the results, it is worth underscoring the limitations of these data: first, as already mentioned, these deviations represent a very loose definition of merit pay (e.g., they may be payments for extra work or cost of living differentials). We are no doubt capturing various kinds of supplemental pay that we

may or may not think of as merit pay. Moreover, since the SASS teacher sample is not designed to be representative for individual districts, our aggregate district measures are suspect and should be taken as no more than suggestive.

Table 5 shows the results from the second-stage regressions.[19] As before, the first column includes the Accountability Index, the second column stratifies the Accountability Index into high and low (the omitted group), the High Accountability measure, and the third column includes the Performance Measure Adjusted Index.

The results in Table 5 provide some evidence that greater performance information may increase informal merit pay in districts, at least as far as we have correctly measured both variables. In column one, the coefficient on the Accountability Index value is positive and significant, indicating that greater degrees of accountability are associated with a $227 increase in deviation from the single salary schedule. Likewise, the coefficient for High Accountability Index in the second regression is positive and significant, adding $435 in salary deviation. The results for the Performance Measure Adjusted Index in the third column are in the expected direction, showing a much smaller $6 gain off the salary schedule, but are not statistically significant.

Interestingly, the figures in Table 5 suggest that the strength of the union influence does not deter unofficial merit pay to the same degree that it deters official merit pay programs. In fact, the coefficients on RTW and the RTW interactions indicate that these states have *smaller* average deviations from the single salary schedule than do non-RTW states. Furthermore, the effect of having no bargaining agreement is negative and significant, indicating that districts with bargaining agreements have larger deviations. This, along with the negative coefficients on the RTW variables, suggests that unofficial and official merit pay may be substitutes for each other.

Alternatively, as noted above, our measures of informal merit pay likely capture other types of compensation that could be correlated with the operating environment of the district.

Table 5: Analysis of Informal Merit Pay
(Standard Errors in Parenthesis)

	(1)	(2)	(3)	(4)	(5)	(6)
Accountability Index	227.0			223.3		
	(46.4)**			(46.1)**		
High Accountability		435.6			436.1	
		(137.8)**			(137.7)**	
Performance Measure Adjusted Index			6.1			5.6
			(26.0)			(26.0)
Right To Work	-854.3	-936.7	-1,040.6	-869.9	-951.0	-1,056.3
	(203.6)**	(201.7)**	(202.0)**	(201.1)**	(200.6)**	(200.7)**
No bargaining agreement	-937.4	-893.6	-872.8	-947.8	-903.0	-882.2
	(236.5)**	(236.3)**	(237.1)**	(235.8)**	(235.9)**	(236.7)**
Meet And Confer	-373.6	-378.7	-460.9	-386.4	-384.3	-467.1
	(361.8)	(363.0)	(362.7)	(361.6)	(362.8)	(362.6)
RTW*No Bargaining	166.9	325.4	460.6	186.9	335.9	472.6
	(326.5)	(324.3)	(325.9)	(325.8)	(323.9)	(325.4)
RTW*Meet And Confer	145.5	276.5	355.9	183.4	286.7	367.3
	(480.9)	(480.3)	(481.7)	(480.1)	(480.0)	(481.4)
Enrollment > 1000	-787.7	-766.0	-720.6	-784.5	-766.8	-721.0
	(173.0)**	(173.1)**	(173.5)**	(172.6)**	(173.1)**	(173.5)**
Avg. School Enrollment/100	52.4	60.2	70.0	50.6	59.1	69.0
	(33.2)	(33.2)	(33.2)*	(33.2)	(33.2)	(33.2)*
Urban District	-338.6	-329.0	-321.4	-348.1	-333.0	-325.2
	(224.0)	(224.2)	(224.9)	(223.8)	(224.1)	(224.8)
Suburban District	364.4	407.8	447.6	366.9	403.6	443.7
	(167.8)*	(167.7)*	(168.3)**	(167.6)*	(167.6)*	(168.2)**
% Minority Enrollment	64.4	163.6	259.0	68.3	162.0	257.9
	(287.0)	(286.4)	(286.0)	(286.9)	(286.3)	(285.9)
% Eligible For Subsidized Lunch	-860.4	-838.0	-841.0	-852.9	-839.0	-842.0
	(325.9)**	(326.3)*	(326.7)*	(325.8)**	(326.3)*	(326.7)**
Log Of Median House Value	-178.0	-198.3	-183.8	-198.0	-205.0	-190.7
	(229.8)	(228.0)	(228.3)	(227.4)	(227.8)	(228.0)
Log Of Median Income	174.4	231.3	211.7	162.6	238.8	219.7
	(421.2)	(421.1)	(421.8)	(420.4)	(421.0)	(421.6)
% Of Pop. 25+ With At Least a BA	651.4	810.9	621.3	844.4	805.8	614.8
	(772.6)	(760.8)	(760.8)	(758.4)	(760.7)	(760.7)
District Offers Formal Merit Pay	-184.3	-155.2	-161.1			
	(231.2)	(231.5)	(231.9)			
Sample Size	4070	4070	4070	4070	4070	4070
R^2	0.05	0.05	0.04	0.05	0.05	0.04

* p < .05, ** p < .01

Cost of Merit Pay

Following the model's assumption about the teacher's risk aversion, our third proposition suggests that average salaries will be higher in districts that use merit pay than in those that do not. Table 6 reports the results of regressions predicting salary as a function of degree, experience, and whether or not a district offers formal merit pay. The specification in column 1 considers a teacher's total income from school (this includes base salary, earnings from extracurricular activities, merit pay or state supplements, and school-based summertime work). Consistent with prior findings about teacher salary schedules, we find that having an MA is worth around $4,000 bump in total income; having a PhD is worth about $8,000 extra dollars in total income from school.[20] Moreover, the formal merit pay coefficient in Table 5 is positive and significant. A teacher's total income in districts with merit pay appears to be about $621 more than in non-merit pay districts.[21] However, it is not entirely clear how to interpret this result because the data do not allow us to separate out the details of how teachers receive merit pay. Some teachers may receive a straightforward merit pay supplement, but others may receive more indirect rewards for an "extra" job (e.g., coaching).

To investigate whether this $621 boost in pay is driven by supplemental pay or by increases in base pay, we run a second specification with base pay as the dependent variable (column 2). The results in column 2 suggest that the increase comes mainly as base pay. Of the $621 bump seen by teachers in merit pay districts (column 1), $465 is in the form of base pay (column 2). One way to interpret this is that, in terms of our formal model, districts may face higher costs in securing teachers' participation (β) than they do in paying out actual marginal performance rewards (α).

Table 6: Estimated Income from School
(Standard Errors in Parenthesis)

	(1)	(2)
District Offers Formal Merit Pay	621.85**	465.62**
	(153.89)	(135.03)
Master's Degree	4,194.37**	4,091.84**
	(100.06)	(87.80)
Phd	7,923.80**	7,345.59**
	(497.80)	(43678)
Experience	920.04**	907.77**
	(17.098)	(15.00)
$Experience^2$	-11.02**	-10.922**
	(0.495)	(0.43)
R^2	0.57	0.62
Sample size	30,139	30,139

* p < .05, ** p < .01

Charter Schools

The above results clearly suggest that politics play an important role in district decisions about formal merit pay. We explore the issue further by looking at a group of schools that are ostensibly free from many institutional restrictions governing traditional public schools, including those that come with collective bargaining. Though state laws differ in detail, public charter schools are independently run schools that receive public funding and are subject to public oversight and accountability but are exempt from many of the rules and regulations that govern traditional public schools, including (often) employment contract rules. For example, in Minnesota, charter schools are not bound by district bargaining agreements, although a charter school's teachers may choose to negotiate with their school as a unit (Education Commission of the States, 2005). In 2000, 28 states had charter school legislation. Three states - Arizona, Michigan, and California - accounted for almost half (47 percent) of the 1,010 charter schools that were in operation that year. Nationally, charter schools enrolled a little over 260,000 students (NCES, 2002).

As mentioned above, states vary in the degree to which charter schools are regulated. As a measure of this, we include rankings developed by the Center for Education Reform (CER), a Washington D.C.-based charter school advocacy group, in our regressions. CER produces a rank-order list of state charter laws from "strong" to "weak," where strength (by their definition) is associated with more flexibility (i.e., *fewer* regulations) for charter schools. So, for example, a charter law that requires or allows charter teachers to be a part of a school district bargaining unit would be considered "weaker" than one that does not. In 2000, CER ranked Arizona as the "strongest" charter law in the country and Mississippi as the "weakest" (Greene, 2000). With this, we hypothesize that the degree of charter regulation will affect the political costs of implementing merit pay. We would expect charters in states with "weak" laws to be less likely to use merit pay.

Charter schools are overall significantly more likely than traditional public schools to use merit pay. For instance, after controlling for the degree of union influence and community demographics, charter schools are 11.5 percentage points more likely to use formal merit pay[22].

Table 7 reports the results of a probit analysis of formal merit-pay use in charter schools. This is equivalent to the probit analysis of public school districts shown in Table 3, except that here the unit of analysis is the charter school, not the district. Also, because so few charter schools have bargaining agreements we were unable to include interactions of RTW and bargaining status.

As was the case with traditional public schools, the estimates for our teacher performance information variables (the Accountability Index, the binary high-low index, and the Performance Measurement Adjusted Index) are not statistically significant, implying that accountability policy does not have an appreciable effect on the probability that a charter school offers merit pay. By contrast, the CER rankings of state charter laws (1 = the strongest; 31 = the weakest) suggest that

charter laws that include more regulation and less flexibility have a strong negative influence on the probability of merit pay. This finding may reflect the fact that charters operating in more regulated environments face higher political costs if they choose to pursue merit pay than schools operating in less regulated environments.

Interestingly, collective bargaining agreements do not appear to have a discernible impact on a charter school's decision to use merit pay. It should be noted, however, that only 19 of the over 800 charter schools in the sample have a collective bargaining agreement. These results, then, should be interpreted with caution.

Table 7: Probability of Merit Pay in Charter Schools

(Standard Errors in Parenthesis)

	PROBIT		
	(1)	(2)	(3)
Accountability Index	-0.014		
	(0.014)		
High Accountability		-0.038	
		(0.041)	
Performance Measure Adjusted Index			-0.013
			(0.009)
Right To Work	-0.036	-0.048	-0.009
	(0.045)	(0.041)	(0.053)
CER charter law rankings	-0.026	-0.026	-0.026
	(0.004)**	(0.004)**	(0.004)**
No bargaining agreement	0.103	0.102	0.101
	(0.114)	(0.115)	(0.115)
Meet and Confer	-0.165	-0.167	-0.164
	(0.112)	(0.111)	(0.112)
Years of Operation	-0.047	-0.047	-0.047
	(0.014)**	(0.014)**	(0.014)**
Urban School	0.072	0.071	0.072
	(0.055)	(0.055)	(0.055)
Suburban School	0.005	0.002	0.003
	(0.058)	(0.058)	(0.058)
% Minority Enrollment	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)
% Eligible for Subsidized Lunch	0.042	0.040	0.043
	(0.062)	(0.062)	(0.062)
Bound by District Salary Agreement	0.060	0.060	0.066
	(0.041)	(0.041)	(0.041)
School enrollment/100	0.028	0.028	0.028
	(0.006)**	(0.006)**	(0.006)**
Log Of Median House Value	-0.044	-0.043	-0.055
	(0.076)	(0.076)	(0.077)
Log Of Median Income	0.046	0.045	0.059
	(0.060)	(0.060)	(0.061)
% Of Pop. 25+ With At Least a BA	-0.202	-0.208	-0.211
	(0.173)	(0.174)	(0.173)
Sample size	846	846	846

* significant at 5%; ** significant at 1%

We also examined informal merit pay in charter schools in an analysis similar to the one we computed for public schools districts. As before, these results should be taken with caution

since the informal merit pay measure - essentially deviations from an assumed salary schedule – is likely measuring things other than merit pay.

Table 8 reports the results of the informal merit pay regressions. The first three include a dummy for whether the school offers official merit pay and the second three omit this control. The results do not change substantively between the two sets of regressions. Furthermore, almost none of the coefficients are statistically significant in any of the regressions. The results suggest that the differences in deviations from the salary schedule in charter schools may be a result of factors such as differences in the cost of living and other factors we are unable to control for in the regressions. Although imperfect, a quick back-of-the-envelope comparison of the informal merit pay measure in charter schools versus public school districts suggests that charter schools resort to smaller deviations from the salary schedule (mean $5,125) than do public school districts (mean $5,490). This may be because charter schools, freed from the union salary constraints, are less likely to resort to informal merit pay.

Table 8. Informal Merit Pay in Charter Schools
(Standard Errors in Parenthesis)

	(1)	(2)	(3)	(4)	(5)	(6)
Accountability Index	101.7			98.0		
	(106.0)			(105.6)		
High Accountability		90.3			81.7	
		(321.3)			(320.3)	
Performance Measure Adjusted Index			174.7			171.2
			(68.2)*			(67.9)*
Right To Work	-347.1	-213.3	-809.5	-349.4	-218.1	-808.1
	(360.3)	(331.7)	(400.7)*	(360.0)	(331.3)	(400.5)*
CER charter law rankings	21.2	22.4	17.2	18.3	20.0	13.4
	(25.1)	(25.4)	(25.0)	(24.3)	(24.5)	(24.2)
No bargaining agreement	563.5	587.5	548.7	579.9	601.7	569.8
	(967.6)	(968.2)	(963.0)	(966.3)	(966.9)	(961.8)
Meet and confer	893.3	947.4	805.1	881.5	936.5	789.1
	(1,037.5)	(1,037.3)	(1,032.3)	(1,036.5)	(1,036.2)	(1,031.4)
Yrs of operation	28.8	29.3	29.9	21.9	23.3	20.8
	(108.9)	(109.0)	(108.5)	(107.7)	(107.8)	(107.2)
Urban school	-135.1	-112.2	-167.4	-132.8	-110.6	-164.5
	(468.9)	(468.6)	(466.6)	(468.5)	(468.2)	(466.3)
Suburban school	-299.2	-280.2	-322.5	-304.4	-285.3	-329.4
	(475.3)	(475.2)	(473.2)	(474.9)	(474.7)	(472.8)
% minority enrollment	0.4	0.7	-0.4	0.6	0.9	-0.1
	(5.4)	(5.5)	(5.4)	(5.4)	(5.4)	(5.4)
% eligible for subsidized lunch	442.3	466.9	384.8	447.5	470.9	391.9
	(519.7)	(519.3)	(517.7)	(519.2)	(518.9)	(517.3)
District salary Agreement	69.4	104.8	-37.2	79.4	113.5	-23.3
	(334.0)	(334.7)	(334.0)	(333.0)	(333.7)	(332.9)
School enrollment/100	3.1	6.0	0.8	6.8	9.2	5.6
	(51.8)	(52.0)	(51.4)	(51.0)	(51.3)	(50.7)
Log Of Median Income	-120.4	-126.4	74.0	-116.5	-122.8	75.5
	(628.3)	(628.7)	(630.5)	(627.9)	(628.2)	(630.2)
Log Of Median House Value	469.9	532.8	204.2	477.2	538.9	216.5
	(495.3)	(496.6)	(503.7)	(494.7)	(496.0)	(503.0)
% Of Pop. 25+ With At Least a BA	-730.2	-811.7	-574.5	-770.0	-847.5	-626.3
	(1,445.2)	(1,453.5)	(1,436.3)	(1,441.4)	(1,449.5)	(1,432.8)
School Offers Formal Merit Pay	137.7	119.8	182.9			
	(315.9)	(315.8)	(314.7)			
R-squared	0.01	0.01	0.02	0.01	0.01	0.02
Sample size	673	673	673	673	673	673

* significant at 5%; ** significant at 1%

Finally, average salaries in charter schools that use merit pay appear higher than in charter schools that do not, about $698 more in total pay – this echoes our findings in public school districts. However, whereas this differential seemed to work through base pay in public school districts, in charter schools it appears to come in the form of additional pay. This is perhaps

unsurprising given that charter schools are more likely to use formal add-on merit pay than public school districts.

Table 9. Estimated Income from School in Charter Schools (Standard Errors in Parenthesis)

	(1)	(2)
District Offers Formal Merit Pay	698.5	371.0
	(326.3)*	(278.1)
Master's Degree	4,822.4	4,410.5
	(351.7)**	(299.8)**
Phd	6,777.4	7,457.0
	(1,503.4)**	(1,281.4)**
Experience	606.6	631.6
	(57.9)**	(49.3)**
$Experience^2$	-7.1	-8.2
	(1.9)**	(1.6)**
R^2	0.46	0.52
Sample size	2189	2189

* $p < .05$, ** $p < .01$

CONCLUSION

Despite the persistence of merit pay as a policy proposal in public education, the theoretical deck appears stacked against it: the "nature of teaching" hypothesis suggests that teaching and performance incentives are a poor fit; the "political cost" hypothesis suggests that merit pay is unlikely because of the power of teacher unions. We present a simple principal-agent model that incorporates both concerns, predicting that districts will be more likely to use merit pay when they have more performance information about teachers and when the influence of teachers unions is weaker. These theoretical predictions, however, are far clearer than the empirical results presented here.

Analyzing district and teacher data from the 2000 SASS, we find some support for the "political cost" hypothesis. As might be expected, collective bargaining agreements have a significant negative effect on the likelihood of districts using merit pay, and state charter laws affect the likelihood of merit pay in charter schools. These findings strongly suggest that

structural reform of the political relationship between school districts and unions would be a necessary condition prior to any widespread use of merit pay as an avenue of school reform.

By contrast, we find little support for the idea that policy changes that increase accountability and sharpen performance measures change the "nature of teaching" in a way that makes formal merit pay more likely. Of course, the failure to find evidence of a relationship between performance information and the formal merit pay decision may, despite our efforts, be due to some unobservable, mediating variable driving results.

ACKNOWLEDGEMENTS

This project is supported by a Field Initiated Study grant form the Office of Educational Research and Department of Educational Research and Improvement, U.S. Department of Education. The views expressed in this paper do not necessarily reflect those of the University of Washington or the U.S. Department of Education. The authors would like to thank Dale Ballou, Larry Kenny, Patrick McEwan, Richard Murnane, seminar participants at the Center on Reinventing Public Education at the University of Washington, and seminar participants at the American Education Finance Association meetings for helpful comments. All errors are solely the responsibility of the authors.

APPENDIX A

In the body of the paper we present the model under the scenario where the district holds no regard for teacher utility and teachers do not care about student achievement. Here we present three alternative scenarios that vary these assumptions, resulting in implications that vary by degree, but not direction.

Scenario II: The district does not care about teachers; teachers care about students

If teachers care about students, their utility becomes a function of their net benefit w and the outcome x.

$$U(w,x) = -\exp\{-r(w+\delta x)\} \tag{17}$$

In (17) δ is the marginal value of the student output x to the teacher. With the linear pay scheme and cost of effort as before,

$$U(w,x) = -\exp[-r\{(\alpha+\delta)x+\beta-\frac{1}{2}t^2\}] \tag{18}$$

The teacher's expected utility becomes

$$E[U(w,x)] = -\exp\{-r(\alpha+\delta)t+\frac{1}{2}r^2\sigma^2(\alpha+\delta)^2-r\beta+\frac{1}{2}rt^2\} \tag{19}$$

And her CE income becomes

$$y = (\alpha+\delta)t-\frac{1}{2}r\sigma^2(\alpha+\delta)^2+\beta-\frac{1}{2}t^2 \tag{20}$$

The teacher's maximization problem becomes

$$\max_t \; y = (\alpha+\delta)t-\frac{1}{2}r\sigma^2(\alpha+\delta)^2+\beta-\frac{1}{2}t^2 \tag{21}$$

The first order condition is $t = \alpha+\delta$. The teacher's maximized CE income under the incentive scheme becomes

$$\tilde{y} = \frac{1}{2}(\alpha+\delta)^2-\frac{1}{2}r\sigma^2(\alpha+\delta)^2+\beta \tag{22}$$

Since the district does not care about the teacher, its maximization problem is the same as it was under Scenario I except for the fact that it now faces a different constraint from the teacher's maximization.

$$\underset{\alpha}{Max}\ (b-\alpha)t-\beta-C=(b-\alpha)(\alpha+\delta)-\beta-C \tag{23}$$

The first order condition is $\alpha=\frac{1}{2}(b-\delta)$. By plugging this into the teacher's certainty equivalent income and the district's net benefit, the district's surplus becomes:

$$NB_{2.1}^{D*}=\frac{1}{4}(b+\delta)^2-C-\beta \tag{24}$$

The maximized joint surplus in this case is:

$$\tilde{S}=\frac{1}{8}(b+\delta)^2(3-r\sigma^2)-C \tag{25}$$

Scenario III: The district cares about teachers; teachers do not care about students

In this case, the district cares about the teacher's utility. As in Scenario I, the teacher does not care about the student achievement outcome and so her maximization problem remains the same. The district's optimal policy is to choose α to maximize the total surplus, i.e., the sum of the teacher's CE income and the district's expected benefit:

$$TS_2=[(b-\alpha)t-C-\beta]+(\alpha t-\frac{1}{2}r\sigma^2\alpha^2+\beta-\frac{1}{2}t^2) \tag{26}$$

Which becomes

$$TS_2=b\alpha-\frac{1}{2}(1+r\sigma^2)\alpha^2-C. \tag{27}$$

The first-order condition for the teacher's maximization is,

$$b=(1+r\sigma^2)\alpha. \tag{28}$$

By substituting the α obtained in equation (28) back into the total surplus equation (27) the maximized total surplus becomes,

$$TS_2^* = \frac{1}{2(1+r\sigma^2)}b^2 - C \tag{29}$$

and,

$$TS_2^* = \frac{1}{2(1+r\sigma^2)}b^2 - C \tag{30}$$

Scenario IV: The district cares about teachers; teachers care about students

Under this scenario, the district's regard for the teacher means that it will seek to maximize joint surplus as it did in Scenario III. The teacher's regard for the student outcomes means that this scenario's utility function is the same as under Scenario II in equation (7) and the maximization problem remains the same.

The district's expected joint surplus becomes $\{(b-\alpha)t - C - \beta\} + \tilde{y}$ or

$$(\alpha+\delta)\{(b-\alpha)+\frac{1}{2}(\alpha+\delta)(1-r\sigma^2)\} - C \tag{31}$$

And the district's maximization problem becomes

$$\max_{\alpha} \mathrm{E}[\tilde{S}] = (\alpha+\delta)\{(b-\alpha)+\frac{1}{2}(\alpha+\delta)(1-r\sigma^2)\} - C \tag{32}$$

The first order condition with respect to α is

$$\alpha = \frac{b-\delta r\sigma^2}{1+r\sigma^2} = b - \frac{br\sigma^2}{1+r\sigma^2} - \frac{\delta r\sigma^2}{1+r\sigma^2} \tag{33}$$

And the maximized joint surplus becomes:

$$\tilde{S} = \frac{(b+\delta)^2}{2(1+r\sigma^2)} - C \tag{34}$$

Table A-1 summarizes the key components of the four scenarios

Table A-1. Models of Four Scenarios for district decisions about merit pay

	α	t	*District Surplus*	*Teacher Surplus*	*Joint Surplus*
Scenario I: The district does not care about teachers; teachers do not care about students.	$\frac{1}{2}b$	$\frac{1}{2}b$	$\frac{1}{4}b^2 - C - \beta$	$\frac{1}{8}b^2(1-r\sigma^2)+\beta$	$\frac{1}{8}b^2(3-r\sigma^2)-C$
Scenario II: The district does not care about teachers; teachers care about students.	$\frac{1}{2}(b-\delta)$	$\frac{1}{2}(b+\delta)$	$\frac{1}{4}(b+\delta)^2 - C - \beta$	$\frac{1}{8}(b+\delta)^2(1-r\sigma^2)+\beta$	$\frac{1}{8}(b+\delta)^2(3-r\sigma^2)-C$
Scenario III: The district cares about teachers; teachers do not care about students	$b-\frac{br\sigma^2}{1+r\sigma^2}$	$b-\frac{br\sigma^2}{1+r\sigma^2}$	$\frac{br\sigma^2}{1+r\sigma^2} - C - \beta$	$\frac{b^2(1-r\sigma^2)}{2(1+r\sigma^2)^2}+\beta$	$\frac{b^2}{2(1+r\sigma^2)}-C$
Scenario IV: The district cares about teachers; teachers care about students.	$b-\frac{br\sigma^2}{1+r\sigma^2}-\frac{\delta r\sigma^2}{1+r\sigma^2}$	$b-\frac{br\sigma^2}{1+r\sigma^2}+\frac{\delta r\sigma^2}{1+r\sigma^2}$	$\frac{(b+\delta)^2 r\sigma^2}{(1+r\sigma^2)^2} - C - \beta$	$\frac{(b+\delta)^2(1-r\sigma^2)}{2(1+r\sigma^2)^2}+\beta$	$\frac{(b+\delta)^2}{2(1+r\sigma^2)}-C$

APPENDIX B:

Percent of Districts with and without collective bargaining in RTW and non-RTW states

	Non-RTW	RTW	Total
No Collective Bargaining	16.34%	23.51%	39.85%
Collective Bargaining	51.1%	9.04%	60.15%
Total	67.44%	32.56%	100%

Count of districts with and without collective bargaining and with and without merit pay in RTW and non-RTW states

	Non-RTW		RTW		Total	
	Merit	*No Merit*	*Merit*	*No Merit*	*Merit*	*No Merit*
No Collective Bargaining	39	482	138	1,072	177	1,554
Collective Bargaining	101	2,097	60	541	161	2,638
Total	140	2,579	198	1,613	338	4192

States and Accountability Indices

State	Carnoy and Loeb Accountability Index	Performance Measure Adjusted Index
Alabama	4	5
Alaska	1	2
Arizona	2	5
Arkansas	1	1
California	4	6
Colorado	1	1
Connecticut	1	4
Delaware	1	1
District of Columbia	5	6
Florida	2	3
Georgia	1	1
Hawaii	1	1
Idaho	2.5	2.5
Illinois	3	8
Indiana	0	0
Iowa	1	3
Kansas	4	7
Kentucky	3	5
Louisiana	1	1
Maine	4	6
Maryland	2	7
Massachusetts	1	1
Michigan	2	2
Minnesota	3	3
Mississippi	1.5	3.5
Missouri	1	1
Montana	0	0
Nebraska	1.5	1.5
Nevada	1	1
New Hampshire	5	6
New Jersey	4	6
New Mexico	5	7
New York	5	10
North Carolina	1	1
North Dakota	3	5
Ohio	1	2
Oklahoma	2.5	4.5
Oregon	1	3
Pennsylvania	1	6
Rhode Island	3	6
South Carolina	1	1
South Dakota	1.5	6.5
Tennessee	5	9
Texas	1	1
Utah	1	3
Vermont	2	3
Virginia	1	2
Washington	3.5	4.5
West Virginia	2	4

REFERENCES

Baker, George. May 2000. "The Use of Performance Measures in Incentive Contracting." *The American Economic Review*, Vol. 90, No. 2, pp. 415-420.

Ballou, Dale. 2001. "Pay for performance in public and private schools." *Economics of Education Review*, Vol. 20, No. 1, pp. 51-61.

The Business Roundtable and National Alliance of Business. 2000. *Pay-for-Performance in Education: An Issue Brief for Business Leaders*. Washington, DC: The Business Roundtable.

Carnoy, Martin, and Susanna Loeb. Winter 2002. "Does External Accountability Affect Student Outcomes? A Cross State Analysis." *Education Evaluation and Policy Analysis,* Vol., 24, No. 4, pp. 305-331.

Clotfelter, Charles T. and Helen F. Ladd. 1996. Recognizing and Rewarding Success in Public Schools, in Helen F. Ladd (ed.), *Holding Schools Accountable: Performance-Based Reform in Education.* Washington, DC: Brookings Institution, pp. 23-64.

Coleman, James S. 1990. *Equality and Achievement in Education* Boulder, CO: Westview Press.

Committee for Economic Development. 2004. *Investing in Learning: School Funding Policies to Foster High Performance: a statement on national policy / by the Research and Policy Committee of the Committee for Economic Development*. Washington, D.C.: Committee For Economic Development.

Dee, Thomas S., and Benjamin Keys. 2004. "Does Merit Pay Reward Good Teachers? Evidence from a Randomized Experiment." *Journal of Policy Analysis and Management*, Vol. 23, No. 3, pp. 471-488.

Dixit, Avinash. 1996. *The Making of the Political Economy: A Transaction-Cost Perspective.* Cambridge, MA: M.I.T. Press.

Dixit, Avinash. Fall 2002. "Incentives and organizations in the public sector: An interpretative review." *Journal of Human Resources*, Vol. 37, No. 4, pp. 696-727.

Education Commission of the States. 2005. *ECS State Policies for Charter School Database*. Online at http://www.ecs.org/html/offsite.asp?document=educationIssues%2FCharterSchools%2FCHDB%5Fintro%2Easp (as of February 16, 2005).

Farkas, Steve, Jean Johnson, and Ann Duffett. 2003. "Stand By Me: What Teachers Really Think About Unions, Merit Pay, and Other Professional Matters." Washington, DC: Public Agenda.

Feldman, Sandra. March 2004. "Rethinking teacher compensation: equitable pay for teachers is part and parcel of improving student performance." *American Teacher*, Vol. 88, No. 6, p 5.

Figlio, David and Lawrence Kenny. 2005. *Do Individual Teacher Incentives Boost Student Performance?* Working Paper

Freeman, Richard and James Medoff. 1984. *What Do Unions Do?* New York: Basic Books.

Goldhaber, Dan. Spring 2002. "Teacher Quality and Teacher Pay Structure: What Do We Know, and What are the Options?" *The Georgetown Public Policy Review* Vol, 7, No. 2, pp. 81-93.

Goldhaber, Dan and Dominic J. Brewer. 1997. "Why Don't Schools and Teachers Seem to Matter? Assessing the Impact of Unobservables on Educational Productivity." *Journal of Human Resources*, 32(3), Summer.

Green, Jay P. 2000. "The Education Freedom Index." *Civic Report No. 14, September 2000*. New York: Center for Civic Innovation, The Manhattan Institute.

Hanushek, Eric A. Summer 1979. "Conceptual and Empirical Issues in the Estimation of Educational Production Functions." *Journal of Human Resources*, Vol. 14, No. 3, pp. 351-88.

Hanushek, Eric A. 1997. "Assessing the Effects of School Resources on Student Performance: An Update." *Educational Evaluation and Policy Analysis*, 19(2), Summer.

Hatry, Harry, John M. Greiner, and Brenda G. Ashford. 1994. *Issues and case studies in teacher incentive plans*. Washington, D.C.: The Urban Institute Press.

Hoff, David J. 2005, February 2. "Governors seek new teacher pay methods: plans include merit pay, bonuses for teaching in struggling schools" *Education Week*, Vol. 24, Issue 21, pp. 22, 28.

Jupp, Brad. 2005, Winter. "The Uniform Salary Schedule." *Education Next*, pp 10-12.

Laffont, Jean-Jacques, and Davit Martimort. 2002. *The theory of incentives: the principal-agent model.* Princeton, N.J.: Princeton University Press.

Lavy, Victor. 2002. "Evaluating The Effect of Teachers' Group Performance Incentives on Pupil Achievement." *Journal of Political Economy*, Vol. 110, No. 6, pp. 1286-1316.

McCaffery, Daniel F., J.R. Lockwood, Daniel M. Koretz, and Laura S. Hamilton. 2003. *Evaluating Value-Added Models for Teacher Accountability*. Santa Monica, Ca: RAND Corporation.

Murnane, Richard J., and David K. Cohen. 1986. "Merit Pay and the Evaluation Problem: Why Most Merit Pay Plans Fail and Few Survive." *Harvard Educational Review*, Vol. 56, No. 1, pp. 1-17.

National Center for Education Statistics. 2002 *Schools and Staffing Survey, 1999-2000: Overview of the Data for Public, Private, Public Charter, and Bureau of Indian Affairs Elementary and Secondary Schools*, NCES 2002-313. Washington, D.C.: U.S. Department of Education.

Odden, Alan, and C. Kelley. 1997. *Paying Teachers for What They Know and Do: New and Smarter Compensation Strategies to Improve Schools,* Thousand Oaks, California: Corwin Press, Inc.

Prendergast, Canice. 1999. "The Provision of Incentives in Firms." *Journal of Economic Literature*, Vol. 37, No. 1, pp. 7-63.

U.S. Department of Education. 1996. "The Patterns of Teacher Compensation." NCES Report 95-829, Washington, DC: National Center for Education Statistics.

Endnotes

[1] See Denver's Joint Taskforce on Teacher Compensation's website http://www.denverteachercompensation.org/ for a full description [Accessed August 4, 2004].
[2] Well-known experimenters Cincinnati, Philadelphia, and Baltimore abandoned their efforts to link pay and evaluation to performance for reasons that both Murnane and Cohen and Ballou would understand: union opposition, the high cost of expanding rewards, and uncertainty about how best to evaluate performance.
[3] We are not interested here in the benefits of merit pay as such, which, in practice, are far from clear. For more research on this topic, see Clotfelter and Ladd (1996), Dee and Keyes (2004), Hatry, Geiner, and Ashford (1994), and Lavy (2002).
[4] Murnane and Cohen (1986) and Ballou (2001) do not develop formal models explaining the decision to offer merit pay.
[5] Of course, districts might adopt other compensation reforms—e.g., pay for knowledge and skills, subject matter specialty, or difficulty of teaching assignment—to accomplish some of the goals mentioned here.
[6] Our model follows along the lines of Dixit (1996) and reaches implications similar to those in Baker (2002).
[7] The model could be extended as in Dixit (1996) to include multiple principals and common agents with similar, though more complex, implications.
[8] Beyond the question of reducing measurement distortion, the σ^2 in equation (12) implies that the probability of utilizing merit pay will rise or fall depending on the degree of influence teachers actually have over student achievement. Since empirical evidence suggests that schooling inputs are more important to disadvantaged students, we might also expect merit pay plans to be more prevalent in districts serving higher proportions of such students, a point we return to later.
[9] In addition to predicting which districts adopt merit pay, the model also makes important predictions about the relative size of merit payments. Unfortunately, we are unable to test this due to data limitations.
[10] Although we would like to use district-level dummies in this first stage estimation, limitations in the data (there are too few teachers per district in the SASS sample) force us to rely on state-level dummies.
[11] The SASS also surveys public school principals, private schools, Bureau of Indian Affairs schools, and public charter schools.
[12] http://nces.ed.gov/surveys/SASS/methods9900.asp
[13] Specifically: 0 = no statewide tests or standards; 1 = statewide test in elementary and middle grades, reporting of results to the state, no sanctions or rewards; 2 = statewide test in elementary and middle grades, moderate accountability sanctions/rewards OR a high school exit test; 3 = statewide test in elementary and middle grades, moderate accountability sanctions/rewards, AND a high school exit test; 4 = statewide tests in elementary and middle grades, strong accountability sanctions/rewards (e.g., threat of reconstitution, loss of students), no high school exit test; 5 = statewide tests in elementary and middle grades, strong accountability sanctions/rewards, AND a high school exit test.
[14] Unfortunately, the SASS 2000 teacher sample does not include district identifiers that allow a direct link from teachers to their districts. To match teachers to districts we had to first match teachers to schools and then match schools to districts. Much of our sample loss was due to teachers that did not match to schools and therefore could not be matched to districts.

[15] Given the SASS sampling strategy, teachers cannot be weighted to be representative of individual districts.

[16] As already noted, statutes in RTW states prevent labor-management agreements that require workers to join a union as condition of employment.

[17] Once again, these results should be interpreted with caution because our accountability variables measure both the level of performance information available *as well as* the performance pressures created by sanctions and rewards. Unfortunately, the data do not allow us to untangle the relative effects of information and performance pressure.

[18] The marginal effects for the interaction terms were computed as

$\frac{\partial^2 \Ö(\cdot)}{\partial x_1 \partial x_2} = \beta_1 x_1 \Ö' + (\beta_1 + \beta_{12} x_2)(\beta_2 + \beta_{12} x_1)\Ö''$ as suggested in Ai and Norton (2003)

[19] The first-stage regressions results were as follows (regressions include state dummies):

Income =	4,446.90*master's degree	8,278.55*PhD	934.44*experience	-11.34*exper^2
	[0.000]	[0.000]	[0.000]	[0.000]

Number in [ ] is p-value n=37,594 R^2= .565

Because some districts in our sample have only a small number of teachers participating in the SASS, we re-estimated the results for only those districts with more than three teachers. However, doing so did not substantively change the coefficients on the variables of interest.

[20] The average pay premiums for advanced degrees are about 11 percent for a Master's degree, 14 percent for an education specialist's degree, and 17 percent for a Doctorate degree (U.S. Department of Education, 1996). The premium for an additional year of service is typically between $1000 and $1500 (Odden and Kelley, 1997).

[21] Table 4 specifications include state dummies, as well as controls for district income and housing value (merging this community-level data causes the loss of some teacher observations).

[22] These differences are based on a probit regression not shown in the paper. These results are available from the authors upon request.