

# The Economics and Politics of Teacher Merit Pay

Andrew Leigh

Parliament of Australia, Canberra ACT 2600, Australia.  
e-mail: [andrew.leigh.mp@aph.gov.au](mailto:andrew.leigh.mp@aph.gov.au)

## Abstract

The debate over merit pay can be summed up as follows: economists like it, voters love it, and teachers are divided. Can merit pay be made to work? I survey three sets of data that are relevant to answering this question: impact studies of teacher merit-pay schemes, evidence on teacher attitudes to merit pay, and surveys of attitudes in the general public to merit pay. Looking at the existing merit-pay plans, one is struck by the fact that their incentive schemes are often very complicated, and most estimates are of short-run effects (so do not capture selection into the teaching profession). Teacher attitudes are mixed, with new teachers more open to merit pay than their more experienced colleagues. Teachers are particularly hostile to merit-pay schemes based on test scores, raising a particular challenge for the political sustainability of such plans. I conclude with 10 suggestions for future research on teacher merit pay.

**Keywords:** Economics of education, teacher quality, performance pay

## 1 Introduction

The debate over merit pay can be summed up as follows: economists like it, voters love it, and teachers are divided. The debate often seems to exude more heat than light: although merit-pay plans have operated since the 19th century, it is only in the past two decades that high-quality academic studies have allowed the debate to move from theory to evidence.

In this article, I survey the available economic studies on the impact of teacher merit pay—with particular reference to the incentive structure embodied in each plan—and discuss what we have learned from the studies to date. I then turn to the political economy of merit pay, compiling evidence on teachers' attitudes towards merit pay. Since teachers have a high degree of autonomy, a merit-pay plan that lacks support from teachers risks being undermined from within. Next, I look at public-opinion surveys on merit pay, and consider what they imply for the political sustainability of merit-pay plans. I conclude with ten suggestions for future research on merit pay.

For the purposes of this article, I define merit pay as encompassing instances in which teachers receive temporary or permanent salary increases for being more effective in the classroom. In this study, I do not regard the following as merit pay: payments for working in hard-to-staff schools, payments for teaching subjects for which there is

an undersupply of applicants, or payments for obtaining additional academic qualifications.

## 2 The Theory of Teacher Merit Pay

The theory of teacher merit pay has been well canvassed elsewhere (see e.g. Lazear 2003; Chevalier et al. 2003; Marsden and Belfield 2006; Podgursky and Springer 2007). In simple terms, merit pay is expected to operate by providing stronger incentives for teachers to work harder, and by encouraging more effective individuals to select into the teaching profession. Lazear (2003) argues that in the case of teacher merit pay, selection effects may be more important than incentive effects. Indeed, if selection effects are positive, then even if teachers are unable to increase their effort level (i.e. incentive effects are zero), merit pay would nonetheless boost student performance.

In making the case for merit pay, there are three pieces of evidence that advocates point towards. First, teachers differ quite markedly in their effectiveness. Averaging across 10 US studies, Hanushek (2011) finds that a 1 standard deviation increase in teacher effectiveness is associated with a 0.13 standard deviation increase in reading scores and a 0.17 increase in math scores. Similar estimates have been found for the variance in teacher effectiveness in the UK (Slater et al. 2009) and Australia (Leigh 2010). One way of expressing these results is that a teacher at the 90th percentile is at least twice as effective as a teacher at the 10th percentile. To the extent that merit-pay schemes are able to induce existing teachers to become more effective, or encourage more capable individuals to choose the profession, they will improve the average level of effectiveness in the teaching profession.

The second argument made in favor of merit pay is that uniform pay scales do not accurately reflect teacher effectiveness. In most developed nations, teachers are paid primarily on experience and academic qualifications. Yet most studies that plot the relationship between teachers' experience and student test score gain show a line that rises in the first few years, and is pretty flat thereafter (e.g. Staiger and Rockoff 2010). Similarly, teachers with a Masters degree do not tend to have higher value-added (e.g. Hanushek and Rivkin 2004, 2006; Leigh 2010).

Third, proponents of merit pay point out that teaching has considerably more pay compression than most occupations. While intra-occupational salary dispersion has risen in most developed nations, it has remained relatively unchanged within teaching (see e.g. Leigh and Ryan 2008). A similar pattern can be seen when following the same individual over time. For example, Chingos and West (2010) show that teachers who are more

effective in the classroom tend to earn more once they leave the teaching profession—despite there being only small pay differences across those who remain in the teaching profession.

Against this, five critiques are commonly made of teacher performance pay.

First, all performance pay schemes are subject to criticism on the basis of unfairness. In the extreme case, suppose that students' performance is only affected by factors outside the school's control. Such factors might include family inputs (e.g. maternal education or the amount of time parents spent reading with the child) or idiosyncratic shocks (e.g. whether a dog is barking outside the classroom on the day of the test). In this scenario, teachers cannot affect students' performance, so the effect of merit pay is merely to introduce random variation into pay rates. To the extent that teachers place a high value on the fairness of their remuneration system, merit pay may end up reducing morale and effort.<sup>1</sup>

Second, test-based merit-pay schemes are criticized on the basis that they may cause teachers to focus on a narrow subset of activities, or to 'teach to the test'. With a perfect test (which took into account not only cognitive skills, but also vital outcomes such as motivation and curiosity), this would not be a problem. In practice, everyone accepts that tests are able to cover only a portion of the desirable outputs, so this remains a live concern.<sup>2</sup> As [Campbell \(1976\)](#) famously notes: 'The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social process it is intended to monitor.' (See also [Rothstein 2009](#); [Neal 2009](#).)

Third, merit-pay schemes that use principal ratings are criticized on the basis that they are vulnerable to favoritism. Particularly in small schools, it is argued, many school principals do not have the skills to administer the pay of the teachers in their schools—and it is better that they be left to focus on educational leadership without having to also worry about salary determination.<sup>3</sup>

Fourth, merit-pay schemes that are based on individual incentives are criticized on the basis that they may reduce the incentive for teachers to collaborate with one another. Compared to some other occupations,

---

<sup>1</sup> For a survey of the literature on fairness in the employment context, see [Bewley \(2004\)](#).

<sup>2</sup> If merit pay is to be based on tests, it is important that those tests cover the desired outcomes as accurately as possible (see e.g. [Hoxby 2002](#), who shows that US states' expenditure on standardised tests is startlingly low).

<sup>3</sup> On the question of whether principals can identify effective teachers, see [Jacob and Lefgren \(2008\)](#). For an analysis of the costs of such an approach, see [Jensen and Reichl \(2011, Appendix B\)](#).

teaching involves relatively little team work.<sup>4</sup> However, there are still instances in which teachers assist one another, through team-teaching, sharing classroom notes, or providing mentoring and advice in the staff-room. Moreover, some experimental evidence suggests that teachers may be more averse to inequality than other professionals (Perez 2010).

Fifth, merit-pay schemes that are based on group incentives (e.g. schemes that provide a reward if the entire school does well) are criticized on the basis that they are vulnerable to a free-rider problem. For example, suppose that teachers in a school of 60 teachers are told that they will all receive merit pay if their Students' test scores exceed a particular threshold. Each teacher will therefore know that her own effort will only make 1/60th of the contribution to the probability that the school meets that threshold.

How should a good merit-pay scheme be designed? In practice, school districts have used a variety of metrics. These include test scores, dropout rates, attendance rates, surveys (of parents, staff and students), principal assessments, and external evaluator reports. From a theoretical standpoint, Neal (2009) argues that merit-pay schemes should include multiple outcome measures, each of which are adjusted for the composition of the student body.

However, as merit-pay schemes include more outcome measures and make more adjustments, they become increasingly complex. As Clotfelter and Ladd (1996: p. 35) note of the performance pay scheme adopted in Dallas in 1991–92:

in its attempt to be scrupulously fair to schools, Dallas has developed an approach that is incomprehensible to most participants in the process and to most outside observers. Even the straightforward form of the two-stage approach . . . is difficult to explain and discuss. The addition of a variety of sophisticated adjustments and refinements may serve to counter various concerns, but in the process it makes the approach even more opaque. Outcome measures that school officials see and understand – such as test scores – go into a black box, and there they are adjusted in various complex ways, standardized, and then restandardized. The indicators that emerge bear little resemblance to the data that originally went in to the black box. School officials neither understand the process nor have any idea what sort of gains would have been required for them to achieve a high ranking.

---

<sup>4</sup> For example, the O\*Net classification of occupations classifies occupations on the question 'How important is it to work with others in a group or team in this job?'. On this basis, elementary school, middle school, and secondary school teachers rank respectively 90th, 451st and 431st out of 861 occupations (see <http://www.onetcenter.org>, last accessed 11 November 2011).

Complexity arises not only from the use of multiple performance measures, but also from the application of adjustment procedures that economists might regard as intuitively sensible are likely to be incomprehensible to non-economists. To parse out the effects of family background, most merit-pay schemes use either value-added scores, residual scores from a regression on socioeconomic variables, or both.

Take for example, the POINT scheme, described in the following terms by Springer et al. (2010: pp. 3–4):

Our value-added measure was based on students' year-to-year growth on TCAP. To control for the possibility that students at different points in the distribution of scores are likely to make different gains, we benchmarked each student's gain against the average gain, statewide, of all students taking the same test with the same prior year score. Benchmarking was simple: we subtracted the statewide average gain from a student's own gain to find out by how much his growth had exceeded the state average. Finally, we averaged these benchmarked scores over a teacher's class—more precisely, over students continuously enrolled in the teacher's class from the twentieth day of the school year to the spring TCAP administration, and for whom we had the prior year scores needed for benchmarking. This average was the value-added score used to determine whether the teacher qualified for a bonus.

This description is likely to produce nods of acceptance from any PhD-trained economist. But given that the typical US teacher is drawn from the bottom half of the aptitude distribution (Corcoran et al. 2004a,b), it would be a mistake to assume that complex incentive schemes are optimal. For example, a school district in central Florida that implemented merit pay based on a 12-item performance matrix found that only about half the teachers understood the process for awarding incentive payments to teachers (Adkins 2004).<sup>5</sup>

As behavioral economics has taught us, it is a mistake to assume that economic actors operate with perfect information. Public finance has shown that many people mistakenly respond to average rather than marginal tax rates. Public choice has shown that ballot order affects election results. These examples suggest that 'cognitive costs' should be incorporated into our standard models of how individuals make decisions. The same is likely to be true of designing merit-pay schemes. If current

---

<sup>5</sup> Forty-five percent disagreed with the statement 'I clearly understand the process for awarding teacher performance pay to individual teachers', whereas 63 percent disagreed with the statement 'I clearly understand the process for calculating the amount of teacher performance pay disbursed to the individual teacher'.

teachers do not understand how a merit-pay plan works, they are unlikely to alter their behavior in the same way that they would if they understood it perfectly. And if potential teachers do not understand merit pay, it is unlikely to affect their selection into the profession.

### 3 Does Merit Pay Work?

In this section, I review the evidence on the impact of merit pay. I focus primarily on test scores, though in cases where researchers have also looked at other outcomes, I also discuss those findings.

Before turning to the available studies, it is worth saying a few words about what the ideal merit-pay study would look like. In my view, it would have the following four characteristics. First, the merit-pay scheme would be one that appeared economically optimal (from a theoretical standpoint). Second, the study would use random assignment, to ensure that measured differences were truly causal (since merit-pay schools might differ in other respects). Third, the study would have a long duration, with randomization carried out across different labor markets (e.g. villages or cities), so that the researcher could estimate both incentive effects (e.g. teacher effort) and selection effects (e.g. more talented people choosing the teaching profession). And fourth, the study would be based on a population comparable to that which the reader is interested in. The last of these criteria is, naturally, in the eye of the beholder. For example, a study that estimated the impact of merit pay in San Francisco schools might be given little weight by a Bangladeshi policymaker.

In [Table 1](#), I summarize the existing studies on merit pay. Mine is not the first attempt to survey this fast-growing literature [see e.g. the excellent surveys of [Podgursky and Springer \(2007\)](#) and [Springer \(2009\)](#)].<sup>6</sup> However, it differs slightly from previous attempts in that I pay special attention to describing the incentives that were at the core of each merit-pay plan. In each case, I describe how the incentive operated, whether it was based on teacher or school performance, and its size relative to annual salary (with reference to external sources where the study does not give the average salary for its teachers). I also provide my own ranking of the simplicity of the incentive program, ranging from 1 (most complex) to 5 (simplest). Studies are given a lower simplicity ranking the more indicators are

---

<sup>6</sup> An earlier version of this article included the Tennessee study of [Dee and Keys \(2004\)](#), which exploited random assignment across teachers in the Project Star experiment. I have opted to omit that study, since it is unclear whether its effects are due to differences in teacher quality or the impact of merit pay.

**Table 1** Impacts of teacher merit pay on student outcomes

Citation	Location	Study period (years when merit pay was in operation)	Incentive level	Incentive basis	Simplicity of incentive (1 = most complex, 5 = simplest)	Incentive size (as percentage of annual salary)	Evaluation methodology	Test score outcomes	Other findings
<b>Randomized evaluations of a single merit pay plan</b>									
Fryer (2011), Marsh et al. (2011)	New York City, NY, USA	2007–2010	School	Composite measure, based on test scores, attendance rates and surveys of students, teachers and parents. Schools are assessed relative to peer schools and all schools in the city, with peer school performance rated three times as highly.	1	Around 4% of salary	Randomized trial	Most measures negative but insignificant	Impact on student absences and behavior negative but insignificant. Treatment teachers no more likely to fill in the survey that comprised part of the incentive scheme.
Springer et al. (2010)	Nashville, TN, USA	2006–2009	Teacher	Math test score gain, relative to rest of state	4	Mean bonus 22% of salary	Randomized trial	Most measures positive but insignificant	Treatment group no more likely to say that merit pay discourages teachers from working together. No evidence of teacher cheating. Treatment group had higher scores in non-incentive subjects (science and social studies). Treatment group teachers exerted more effort. No evidence of teacher cheating.
Muralidaran and Sundararaman (2011a), Muralidaran (2011)	Andhra Pradesh, India	2005–2010	Teacher and school	Bonus = Rs 500 * (% Gain in average test scores – 5%) if Gain > 5%	5	Mean bonus 3% of salary	Randomized trial	After 5 years, teacher incentives: math scores 0.54 SD higher and language higher. School incentives: math scores positive but insignificant, language scores 0.14 SD higher.	

(continued)

**Table 1** Continued

Citation	Location	Study period (years when merit pay was in operation)	Incentive level	Incentive basis	Simplicity of incentive (1 = most complex, 5 = simplest)	Incentive size (as percentage of annual salary)	Evaluation methodology	Test score outcomes	Other findings
Glewwe et al. (2010)	Busia and Teso districts, Kenya	1998–1999	School	If school's average test scores in grades 4–8 met certain benchmarks (absolute and improvement), then a teacher in that school had a chance of winning a prize	4	Prizes were in-kind (eg, suit, tea set, bed linens), and were worth 2–4% of salary.	Randomized trial	Scores on merit pay test rose 0.22 SD. Impacts not sustained the year after merit pay ceased, and no significant impact on other tests.	Little impact on teacher attendance. No evidence that the incentive affected teacher cheating.
<b>Natural experiment evaluations of a single merit-pay plan</b>									
Winters et al. (2007)	Little Rock, AK, USA	2003–2006	Teacher	Average test score gain	4	Up to 28% of salary	Matched differences-in-differences (three control schools were selected on the basis of having similar demographic and achievement at baseline).	0.17 to 0.2 SD improvement in math scores	
Lavy (2009)	Israel	2000–2001	Teacher	Tournament based on average student performance relative to predicted, with prediction based on socioeconomic status plus grade and school fixed effects.	4	6–25% of salary per class	Regression discontinuity (based on measurement error in the cutoff score)	Math score 5.3 points higher. English score 2.5 points higher.	Teachers in treatment schools exerted more effort (primarily towards weaker students), were more likely to use small group teaching, individualized instruction and ability tracking, and adapted teaching methods to students' ability. No evidence of teacher cheating.

(continued)



**Table 1 Continued**

Citation	Location	Study period (years when merit pay was in operation)	Incentive level	Incentive basis	Simplicity of incentive (1 = most complex, 5 = simplest)	Incentive size (as percentage of annual salary)	Evaluation methodology	Test score outcomes	Other findings
Lavy (2002)	Israel	1996-1997	School	Tournament with prizes paid if the school was in the top third for improvement on three metrics (average number of credit units per student, the proportion of students receiving a matriculation certificate, and the school dropout rate).	1	1-3% of salary	Regression discontinuity (eligible schools were one of their kind in the community; so Lavy uses as controls the schools that were two of their kind in the community)	Test scores around 0.1 SD higher	Program also had significant effects on reducing the dropout rate and increasing the number of credits taken.
Atkinson et al. (2009)	England, UK	1999-2002	Teacher	Experienced teachers could apply to pass the Performance Threshold. This was judged on five areas: knowledge and understanding of teaching, teaching management and assessment, wider professional effectiveness, professional characteristics, and pupil progress	2	9% of salary (permanently)	Difference-in-difference (control group is less experienced teachers, with controls for experience)	Value-added test scores 0.7 SD higher	

(continued)

**Table 1 Continued**

Citation	Location	Study period (years when merit pay was in operation)	Incentive level	Incentive basis	Simplicity of incentive (1 = most complex, 5 = simplest)	Incentive size (as percentage of annual salary)	Evaluation methodology	Test score outcomes	Other findings
Eberts et al. (2002)	Michigan, USA	1996–1999	Teacher	Share of students completing the course, plus results from student evaluations	3	Up to 20% of salary	Matched differences-in-differences (control school chosen based on conversations with local educators)	Negative but insignificant	Improvement in course completion rate (directly tied to incentive), but deterioration in daily attendance rate (not directly tied to incentive)
Vigdor (2008)	North Carolina, USA	1996–2007	School	Test score gains relative to state average (adjusted for mean-reversion)	4	2 or 4% of salary	Regression discontinuity	Failure to receive a bonus spurs test score gains of 0.02 to 0.03 SD	
Ladd (1999)	Dallas, TX, USA	1991–1995	School	Tournament based on test score gains on two tests (residuals from a regression on socioeconomic variables), plus other variables such as attendance and dropout	2	3% of salary	Difference-in-difference (control group is five other large US cities)	0.2 to 0.4 SD improvement in math and reading	Improvements in dropout and attendance. Higher rates of principal turnover

**Comparisons across multiple jurisdictions with differing types of merit pay**

Woessmann (2011)	28 OECD nations	2003	Mixed	Mixed	Mixed	Mixed	OLS with controls	PISA math score 0.24 SD higher	
Figlio and Kenny (2007)	USA	1992	Teacher	Mixed	Mixed	Bonus at least 10% of salary, or given to less than 20% of teachers	OLS with controls	NELS 12th grade scores (sum of reading, math, science and history) 0.04 to 0.06 SD higher	

OLS: Ordinary least squares; NELS: National Educational Longitudinal Survey; PISA: Programme for International Student Assessment.

used to make up the merit rating, and the more adjustments that are made to the data.<sup>7</sup> Naturally, such a ranking is subjective.

Since the best causal impacts are those estimated from random experiments, I place these studies into a separate group at the top of the table. These are followed by natural experiment studies of particular merit-pay plans, and comparisons across multiple merit-pay plans in different jurisdictions.

Of the four random assignment studies, only two presented teachers with incentives that were both simple and certain. In the Indian experiment of [Muralidaran and Sundararaman \(2011a\)](#) and [Muralidaran \(2011\)](#) and in the Nashville study of [Springer et al. \(2010\)](#), teachers were paid a bonus that was a direct function of test score growth. In contrast, the New York City experiment analyzed by [Fryer \(2011\)](#) and [Marsh et al. \(2011\)](#) had quite a complex incentive structure. In the New York experiment, there were a large number of inputs into the incentive formula, which were then weighted three times as heavily for peer schools as for other schools in the city. In the Kenya experiment of [Glewwe et al. \(2010\)](#), the incentive criteria were relatively simple, but payment was uncertain, with merit effectively rewarded with a ticket in a lottery whose prizes are in-kind rather than cash.

Yet this distinction does not map perfectly onto the random assignment studies. Of the two studies that I would regard as both simple and certain, the Indian experiment had a positive and significant impact on test scores, while the Nashville study did not. Among those that were either complex or uncertain, the New York City study had no impact, and the Kenyan study had an initial effect which was not sustained after the program ceased.

In the next panel, I set out seven natural experiment studies of particular merit-pay programs. These studies use two approaches to determine the appropriate counterfactual: regression discontinuity and differences-in-differences. While determining the appropriate comparison group is always difficult, I am inclined to give somewhat less weight to two studies. [Atkinson et al. \(2009\)](#) uses less-experienced teachers as a control for more experienced teachers (and is therefore sensitive to the particular parameterization of experience). In the case of [Eberts et al. \(2002\)](#), my concern arises because the study is based on only one treatment school, with the

---

<sup>7</sup> As noted above, there are multiple ways that merit-pay schemes parse out family background. For example, some schemes use value-added scores (e.g. [Springer et al. 2010](#); [Muralidaran and Sundararaman 2011a](#)), others use regression-adjustment (e.g. [Lavy 2009](#)), and some use both regression-adjustment and value-added scores (e.g. [Ladd 1999](#)). Another approach is to use changes in school averages from year-to-year (e.g. [Glewwe et al. 2010](#); [Fryer 2011](#)).

choice of control school based on conversations with local educators (both schools are therefore sensitive to idiosyncratic shocks).

Atkinson et al. (2009) estimates an extremely large treatment effect (0.7 standard deviations), whereas Eberts et al. (2002) find effects that are negative and statistically insignificant. Of the other five studies, all find positive impacts of merit pay, with some weak support for the notion that simpler merit-pay schemes produce larger student test score gains. For example, the Little Rock merit-pay scheme described by Winters et al. (2007) was relatively simple, operated at a teacher level, provided large bonuses (up to 28% of salary) and produced test score gains of up to 0.2 standard deviations. It is also worth noting that the analysis of North Carolina's merit-pay plan (Vigdor 2008) covers over a decade, and is therefore likely to have captured both selection and incentive effects.<sup>8</sup>

Finally, I set out the results from two studies that compare across merit-pay plans. Both across school districts (Figlio and Kenny 2007) and across Organisation for Economic Co-operation and Development (OECD) countries (Woessmann 2011), merit pay appears to be associated with higher student test scores. Although these studies do not permit us to say much about the impact of incentive size or simplicity, they do strongly suggest that the presence of merit pay is associated with higher test scores.

The longer that I gaze at Table 1, the more I come to the view that we have more 'theories' of merit pay than we have good merit-pay 'studies'. If we are to unpack the impact of the incentive level (teacher or school), incentive size, and the simplicity of the scheme, we need more merit-pay experiments. Studies that compare two or more different merit-pay systems will be particularly useful. For example, the Indian studies of Muralidaran and Sundararaman (2011a) and Muralidaran (2011) show that while school-level and teacher-level merit pay both raise test scores, teacher-level merit pay is considerably more effective.

Another issue worth studying is treatment heterogeneity. For example, Muralidaran and Sundararaman (2011a) find little evidence of treatment heterogeneity by baseline test scores, student characteristics, and school characteristics; though there is some evidence that better-educated teachers respond more strongly to incentive payments. Similarly, Fryer (2011) finds little evidence of treatment heterogeneity in New York City schools that are eligible for school-based merit pay.

Several studies also look at potential adverse impacts of merit pay, and are generally able to reject claims that merit pay has negative side-effects.

---

<sup>8</sup> Unfortunately, because the North Carolina plan was implemented statewide, Vigdor (2008) is unable to estimate the effect of being subject to merit pay. Instead, his regression discontinuity estimate looks at the effect that narrowly missing out on a merit bonus has on a school's gains the following year.

On ‘teaching to the test’, Muralidaran and Sundararaman (2011a) find that the treatment group had higher scores in non-incentive subjects. On collegiality, Springer et al. (2010) find no evidence that merit pay discourages teachers from working together. On discouragement effects, Lavy (2009) and Muralidaran and Sundararaman (2011a) find that merit pay increases teacher effort. And on cheating, Lavy (2009), Glewwe et al. (2010), Springer et al. (2010), and Muralidaran and Sundararaman (2011a) find no evidence that merit pay increases teacher cheating.

It is also worth recognizing the fact that other more popular educational interventions may be less effective or more expensive than merit pay. For example, Hanushek (1998), Hoxby (2000), and Rockoff (2009) find zero or negligible benefits of across-the-board class size reductions. An oft-cited upper bound estimate of the impact of reducing class sizes is Krueger (1999), who finds that a one-sixth reduction in class size raises test scores by 0.22 standard deviations. The same expenditure would facilitate a merit-pay plan that paid 50% salary bonuses to 1 in 12 teachers.

#### 4 What Do Teachers Think About Merit Pay?

In Table 2, I outline results from a number of surveys of teachers about their attitudes to merit pay.<sup>9</sup> Although the typical survey finds that a majority of teachers are opposed to merit pay, survey results tend to be quite sensitive to the wording of the question. For example, a 1983 US survey found that 63% of teachers supported larger salary increases for more ‘effective’ colleagues, while a poll the next year found 64% of teachers opposed to ‘merit pay’. In 1987–1988, a US survey found that 71% of teachers supported promotion based on performance, 55% supported merit bonuses, and 66% supported school-wide bonuses.

However, when US teachers are asked in more detail how they think merit pay should operate, it emerges that teachers most dislike merit-pay schemes that rely on standardized tests. In an extensive 2007 survey, merit pay based on principal evaluations received 58% support, while merit pay based on test scores received 34% support. Similarly, in a 2006 study of teachers in Washington state, 83% opposed merit pay based on test score gains.

Probing further about the concerns that teachers have about merit pay, a 2003 survey uncovered the fact that 63% of teachers think merit pay would be more likely to lead to competition/jealousy than an improvement

---

<sup>9</sup> In this article, I do not explore principals’ attitudes to merit pay, due to a paucity of such surveys. One Australian survey suggests that principals are about as lukewarm as teachers towards merit pay that is test-score based (McKenzie et al. 2008).

Table 2 Teacher attitudes to merit pay

Country/ region	Year	Subject pool	Question	Responses	Citation
USA	1983	All teachers	Teachers who are more effective in the classroom should receive larger salary increases than teachers who are less effective	63% agree, 37% disagree (sample size: 1,261)	Rist (1983)
USA	1984	All teachers	How do you, yourself, feel about the idea of merit pay for teachers? In general do you favor or oppose it?	32% favor, 64% oppose, 4% no opinion (sample size: 813)	Elam (1989: p. 790)
USA	1989	All teachers	How do you, yourself, feel about the idea of merit pay for teachers? In general do you favor or oppose it?	31% favor, 61% oppose, 8% no opinion (sample size: 830)	Elam (1989: p. 790)
USA	1987– 1988	All teachers	For each of the following pay incentives, please indicate whether you favor or oppose the incentive... Salary increases as part of a career ladder in which teachers progress through several promotional levels based on their performance	41% strongly favor, 30% mildly favor, 12% mildly oppose, 17% strongly oppose (sample size: 33,865)	Ballou and Podgursky (1993)
USA	1987– 1988	All teachers	For each of the following pay incentives, please indicate whether you favor or oppose the incentive... A merit pay bonus for exceptional performance in a given year	29% strongly favor, 26% mildly favor, 16% mildly oppose, 29% strongly oppose (sample size: 33,865)	Ballou and Podgursky (1993)
USA	1987– 1988	All teachers	For each of the following pay incentives, please indicate whether you favor or oppose the incentive... A schoolwide bonus for all teachers in a school that shows exceptional performance or improvement in a given year	35% strongly favor, 31% mildly favor, 14% mildly oppose, 21% strongly oppose (sample size: 33,865)	Ballou and Podgursky (1993)

(continued)

**Table 2 Continued**

Country/ region	Year	Subject pool	Question	Responses	Citation
USA	2003	All teachers	How much would you favor or oppose giving financial incentives to ... teachers who consistently receive outstanding evaluations by their principals	28% strongly favor, 34% somewhat favor, 13% somewhat oppose, 20% strongly oppose, 6% not sure (sample size: 1,345)	Farkas et al. (2003)
USA	2007	All teachers	How much would you favor or oppose giving financial incentives to ... teachers who consistently receive outstanding evaluations by their principals	24% strongly favor, 34% somewhat favor, 18% somewhat oppose, 21% strongly oppose, 3% not sure (sample size: 1,010)	Duffett et al. (2008)
USA	2003	All teachers	How much would you favor or oppose giving financial incentives to ... teachers who consistently work harder, putting in more time and effort than other teachers	31% strongly favor, 36% somewhat favor, 14% somewhat oppose, 15% strongly oppose, 4% not sure (sample size: 1,345)	Farkas et al. (2003)
USA	2003	All teachers	How much would you favor or oppose giving financial incentives to ... teachers whose kids routinely score higher than similar students on standardized tests	12% strongly favor, 27% somewhat favor, 24% somewhat oppose, 32% strongly oppose, 6% not sure (sample size: 1,345)	Farkas et al. (2003)
USA	2007	All teachers	How much would you favor or oppose giving financial incentives to ... teachers whose kids routinely score higher than similar students on standardized tests	11% strongly favor, 23% somewhat favor, 25% somewhat oppose, 39% strongly oppose, 3% not sure (sample size: 1,010)	Duffett et al. (2008)

(continued)

**Table 2 Continued**

Country/ region	Year	Subject pool	Question	Responses	Citation
USA	2003	All teachers	Suppose that in your district the students of some teachers make more academic progress - in terms of improved reading levels, teacher evaluations and classroom tests - when compared to similar students taught by other teachers. How much would you favor or oppose financially rewarding those teachers?	Strongly favor 12%, somewhat favor 36%, somewhat oppose 20%, strongly oppose 25%, not sure 8% (sample size: 1,345)	Farkas et al. (2003)
USA	2007	All teachers	Suppose that in your district the students of some teachers make more academic progress - in terms of improved reading levels, teacher evaluations and classroom tests - when compared to similar students taught by other teachers. How much would you favor or oppose financially rewarding those teachers?	Strongly favor 10%, somewhat favor 34%, somewhat oppose 22%, strongly oppose 29%, not sure 5% (sample size: 1,010)	Duffett et al. (2008)
USA	2003	All teachers	Which comes closer to your view: Teachers can and do make a difference in what kids learn and they should be financially rewarded when they succeed OR It's not fair to hold teachers accountable when so many things that affect student learning are beyond their control	29% rewarded, 59% not fair, 12% not sure (sample size: 1,345)	Farkas et al. (2003)

(continued)



**Table 2 Continued**

Country/ region	Year	Subject pool	Question	Responses	Citation
USA	2003	All teachers	If some form of merit pay for teachers was implemented at your school, which do you think would be more likely to happen: It would give principals a way to reward the teachers who really help kids learn OR Principals would play favorites and reward teachers who are loyal to them or who don't rock the boat	23% principal's reward, 52% principal's play favorites, 25% not sure (sample size: 1,345)	Farkas et al. (2003)
USA	2003	All teachers	If some form of merit pay for teachers was implemented at your school, which do you think would be more likely to happen: Instead of cooperation, there would be unhealthy competition and jealousy among teachers OR Teachers would be motivated to work harder and find ways to be more effective	63% jealousy, 22% more effective, 15% not sure (sample size: 1,345)	Farkas et al. (2003)
USA	2003	All teachers	Do you think that implementing merit pay would make the teaching profession more appealing to the best and brightest, or would it not have that effect?	27% would make it more appealing, 53% would not have that effect, 20% not sure (sample size: 1,345)	Farkas et al. (2003)
Washington state, USA	2006	All teachers	How much do you favor or oppose giving extra compensation to the following types of teachers?... teachers whose students make greater gains on standardized tests than similar students taught by other teachers	Strongly favour 3%, Somewhat favor 14%, Somewhat oppose 23%, Strongly oppose 60% (sample size: 3,121)	Goldhaber et al. (2011)

(continued)

Table 2 Continued

Country/ region	Year	Subject pool	Question	Responses	Citation
Hillsborough County Florida, USA	2008	All teachers	Incentive pay for teachers based on overall performance at the school is a positive change to teacher pay practices	Agree strongly 13%, Agree 25%, Disagree 27%, Disagree strongly 32%, DK 3% (sample size: 1,691)	Jacob and Springer (2008)
Hillsborough County Florida, USA	2008	All teachers	Incentive pay for teachers based on group performance (i.e., gradelevel, department, or interdisciplinary team) is a positive change to teacher pay practices	Agree strongly 5%, Agree 17%, Disagree 33%, Disagree strongly 41%, DK 3% (sample size: 1,691)	Jacob and Springer (2008)
Hillsborough County Florida, USA	2008	All teachers	Incentive pay for teachers based on individual teacher performance is a positive change to teacher pay practices	Agree strongly 22%, Agree 27%, Disagree 20%, Disagree strongly 28%, DK 3% (sample size: 1,691)	Jacob and Springer (2008)
Australia	2007	Primary teachers	To what extent do you agree that the following would help retain teachers in the profession? ... Higher pay for teachers whose students achieve specified goals	Strongly agree 13%, Agree 11%, Disagree 34%, Strongly disagree 37%, Unsure/missing 4% (sample size: 5,209)	McKenzie et al. (2008: p. 98)
Australia	2007	Secondary teachers	To what extent do you agree that the following would help retain teachers in the profession? ... Higher pay for teachers whose students achieve specified goals	Strongly agree 13%, Agree 14%, Disagree 35%, Strongly disagree 32%, Unsure/missing 5% (sample size: 5,394)	McKenzie et al. (2008: p. 99)

(continued)

**Table 2** Continued

Country/ region	Year	Subject pool	Question	Responses	Citation
Andhra Pradesh, India	2006– 2007	All teachers	What is your overall opinion about the idea of providing high-performing teachers with bonus payments on the basis of objective measures of student performance improvement?	51% very favorable, 35% somewhat favorable, 8% neutral, 5% somewhat unfavorable, 2% very unfavorable (sample size: 1,512)	Muralidaran and Sundararaman (2011b)
UK	2000	All teachers	The principle of relating teachers' pay to performance is a good one	Agree 24%, neutral 12%, disagree 63% (sample size: Just under 3000)	Marsden (2000)
UK	2000	All teachers	The principle that individual teachers' pay should take some account of pupil progress is a good one	Agree 27%, neutral 17%, disagree 56% (sample size: Just under 3000)	Marsden (2000)

Surveys are ordered by region and year, except that similarly-worded questions are grouped together.

in effectiveness; 59% say that it's not fair to hold teachers accountable for student learning; and 52% are concerned that merit pay would cause principals to play favorites.

Other research has looked at attitudes to merit pay across different groups of teachers. Merit pay is more likely to find favor among teachers who are less experienced (Ballou and Podgursky 1993; Farkas et al. 2003; Goldhaber et al. 2011). According to Ballou and Podgursky (1993) and Goldhaber et al. (2011), male teachers are more likely to support merit pay than female teachers, while black and Hispanic teachers are substantially more likely to support merit pay than white teachers. Importantly, Perez et al. (2011) find that more effective teachers are more likely to prefer merit pay (a one standard deviation increase in effectiveness correlates with a 3 percentage point increase in the probability of favoring competitive bonuses over uniform salary increases). Further research on this point would be valuable.

There is little evidence that being subject to merit pay has an impact on teachers' attitudes towards it. Ballou and Podgursky (1993) find that teachers in school districts with merit-pay plans are no more hostile to merit pay. Similarly, Springer et al. (2010) find: 'A majority (64%) agreed with the statement: "Teachers should receive additional compensation if their students show outstanding achievement gains" in spring of 2007. Two years later this figure was virtually unchanged (66%). There were no significant differences across grades or between treatment and control groups'. Witham et al. (2008) find an improvement in teacher attitudes to merit pay after teachers were exposed to the program.

Attitudes in other countries vary markedly. In Australia, where merit pay is extremely rare, only 24% of primary teachers (and 27% of secondary teachers) agree that merit pay would help retain people in the profession. In the UK, a 2000 survey (immediately preceding the introduction of performance pay in that country) found a majority of teachers disagreed with the general principle of performance-related pay, and disagreed with the specific notion that teacher pay should take some account of pupils' progress.

In contrast, Muralidaran and Sundararaman (2011b) found that 86% of teachers in the Indian state of Andhra Pradesh held a favorable view of merit pay. Consistent with the findings of Springer et al. (2010), support for merit pay was 5 percentage points higher where teachers were randomly assigned to a school-based merit-pay scheme, and 8 percentage points higher in schools with individual-based merit pay (bonuses had not been distributed at the time of the survey).

Overall, I draw three lessons from surveying teacher attitudes to merit pay. First, teachers have real concerns about the impact of merit pay, particularly on the cooperative spirit in the school. Second, there is

significant heterogeneity in attitudes, with new hires and minority teachers substantially more inclined to support merit pay. Third, the kind of merit-pay plan that receives least support is one that uses test scores.

How can merit pay advocates address these issues? On the first issue, it is important for future merit-pay experiments to look carefully at the impact on collegiality. Some existing studies have taken this approach (e.g. [Springer et al. 2010](#) found that treatment group were no more likely to say that merit pay discourages teachers from working together), but since it is one of the leading objections from educators, future studies should look carefully at it.

The second issue—substantial heterogeneity in attitudes—is one likely explanation as to why opt-in merit-pay schemes (or merit-pay schemes that are only compulsory for new hires) appear to be more politically sustainable than schemes that force veteran teachers to join. For example, some have suggested a ‘grand bargain’ over merit pay, in which teachers are offered the option of choosing a new contract that provides higher risk and higher expected return—but are also free to stay on their existing contracts if they so wish.<sup>10</sup> Such an approach would still encourage more talented people to enter the teaching profession.

The third issue is the trickiest for economists, since it suggests that the simplest and most objective measure of teacher quality (test scores) is also the measure that teachers would least like to be measured on. This suggests that so long as merit-pay schemes are based on test scores, a ‘grand battle’ over merit pay is likely to persist for some time.

## 5 What Do Voters Think About Merit Pay?

Finally, I survey the available public opinion about merit pay ([Table 3](#)). Asked whether teachers should be paid based on the quality of their work or on a standard scale, 71% of US respondents in 2010 supported merit pay—up from 61% in 1983. In another 2010 survey, US respondents were asked whether teacher pay should be tied to students’ academic achievement. Seventy-three percent said that pay should be linked to performance very closely or somewhat closely (up from 60% in 2000). Asked specifically about the use of tests to measure teacher quality, a 2010 survey found 64% of the US public in favor—up from 49% in 2004 (though the wording differed slightly). In a 2011 survey that allowed a ‘neither favor nor oppose’ option, 47% still favored merit pay (a similar figure over the period 2007–2011). Outside the USA, evidence on public attitudes to

---

<sup>10</sup> See [Nadler and Wiswall \(2009\)](#) for a model of how risk aversion affects teachers’ propensity to opt in to merit pay.

**Table 3** Attitudes of the general public to merit pay

Country/ region	Year	Question	Responses	Citation
USA	1983	Should each teacher be paid on the basis of the quality of his or her work, or should all teachers be paid on a standard-scale basis?	Work quality 61%, standard scale 31%, DK 8% (sample size: 1,540)	Gallup (1983)
USA	2010	Should each teacher be paid on the basis of the quality of his or her work, or should all teachers be paid on a standard-scale basis?	Work quality 71%, standard scale 27%, DK 2% (sample size: 1,008)	Bushaw and Lopez (2010)
USA	2000	In your opinion, how closely should a teacher's salary be tied to his or her students' academic achievement: very closely tied, somewhat closely tied, not very closely tied, or not at all tied?	Very closely 25%, somewhat closely 35%, not very closely 12%, not at all 24%, DK 4% (sample size: 1,093)	Rose and Gallup (2000)
USA	2010	In your opinion, how closely should a teacher's salary be tied to his or her students' academic achievement: very closely tied, somewhat closely tied, not very closely tied, or not at all tied?	Very closely 19%, somewhat closely 54%, not very closely 14%, not at all 13%, DK < 1% (sample size: 1,008)	Bushaw and Lopez (2010)

(continued)

Table 3 Continued

Country/ region	Year	Question	Responses	Citation
USA	2004	In your opinion, should one of the measurements of a teacher's quality be based on how well his or her students perform on standardized tests, or not?	Yes 49%, No 47%, DK 4% (sample size: 1,003)	Crabtree (2004)
USA	2010	Should teachers' evaluations be based in part on their students' progress on standardized tests?	Yes 64%, No 31%, DK 4% (sample size: 1,000)	Time Magazine (2010)
USA	2007	Do you favor or oppose basing a teacher's salary, in part, on students' academic progress [on state tests]?	Completely favor 14%, somewhat favor 31%, neither favor nor oppose 24%, somewhat oppose 16%, completely oppose 15% (sample size: 2,000)	Howell et al. (2007)
USA	2008	Do you favor or oppose basing a teacher's salary, in part, on his or her students' academic progress on state tests?	Completely favor 13%, somewhat favor 31%, neither favor nor oppose 28%, somewhat oppose 16%, completely oppose 12% (sample size: 1,608)	Howell et al. (2008)
USA	2009	Do you favor or oppose basing a teacher's salary, in part, on his or her students' academic progress on state tests?	Completely favor 13%, somewhat favor 30%, neither favor nor oppose 30%, somewhat oppose 15%, completely oppose 12% (sample size: 3,251)	Howell et al. (2009)

(continued)

Table 3 Continued

Country/ region	Year	Question	Responses	Citation
USA	2010	Do you favor or oppose basing a teacher's salary, in part, on his or her students' academic progress on state tests?	Completely favor 15%, somewhat favor 34%, neither favor nor oppose 25%, somewhat oppose 15%, completely oppose 11% (sample size: 2,776)	Howell et al. (2010)
USA	2011	Do you favor or oppose basing the salaries of teachers around the nation, in part, on their students' academic progress on state tests?	Completely favor 14%, somewhat favor 33%, neither favor nor oppose 26%, somewhat oppose 18%, completely oppose 9% (sample size: 2,600)	Howell et al. (2011)
USA	2010	Do you support or oppose 'merit pay', the concept of paying teachers according to their effectiveness?	Yes 71%, No 23%, DK 6% (sample size: 1,000)	Time Magazine (2010)
Quebec, Canada	2011	Teacher remuneration is currently based on seniority and level of education. Do you believe that teacher performance should be evaluated and that the best teachers should be better paid?	Yes 67%, No 23%, DK 10% (sample size: 1,005)	Leger Marketing (2011)

Surveys are ordered by region and year, except that similarly-worded questions are grouped together.



merit pay is scant, though a 2011 survey in Quebec found two-thirds in favor of evaluating teacher performance and paying the best teachers more.

The growing support for merit pay in the USA most likely reflects the fact that merit-pay reforms have received political support from both major parties in recent years (e.g. in the 2008 Presidential election, both John McCain and Barack Obama advocated more merit pay). Indeed, a survey that explicitly reminded voters of President Obama's support for merit pay boosted support for the policy among the US public by 13 percentage points (Howell et al. 2009). In addition, growing support for merit pay may also be due to a rise in earnings inequality (and perhaps more merit pay) in the non-teaching labor market.

However, political support for merit pay does not necessarily translate into political sustainability. When budgets are tight, a merit-pay program that provides year-to-year salary bonuses is an easy candidate for abolition. This may explain why 75% of merit-pay programs that one set of researchers observed in 1983 were defunct a decade later (Hatry et al. 1994, cited in Eberts et al. 2002).

There are a myriad of ways to block and dilute merit-pay plans, as Buck and Greene (2010) expertly document. For example, they point to an Arizona school district which set its performance target as teaching students 10% more algebra than they knew before taking the class. Given that the average improvement was 90%, this was not a difficult threshold. Another example of dilution is the Denver ProComp scheme, under which the largest bonus payment (a permanent \$3300 increase in salary) is for attaining a graduate degree.

## 6 A Research Agenda for Merit Pay

In Section 3, I set out the desirable characteristics of a merit-pay study: an economically optimal merit-pay design, random assignment, long duration, variation across labor markets, and a context sufficiently similar for the reader to feel that it is relevant to them. Unsurprisingly, none of the existing studies satisfy all of these criteria. From a theoretical standpoint, the incentive criteria in several merit-pay studies appear opaque or overly complicated. If, as the old saying goes, a camel is a horse designed by a committee, then the merit-pay world is replete with camels.

From a political economy standpoint, it is easy to see why overly complex schemes emerge. In balancing competing interests, there will often be a strong temptation to include another indicator to appease a particular sectional interest. In a roomful of experts, there will be a tendency towards measures that are complicated and accurate (to use an econometric

analogy, there is no penalty function for adding additional parameters to the model in order to improve goodness-of-fit). In my experience, it is much harder to be the person who argues for an outcome measure that is simpler and less accurate. The result is generally a merit-pay plan that is hard to explain to teachers and the general public.

Yet there is clearly strong public support for merit pay, and enough evidence from the existing economic studies to support further merit-pay trials. Anyone who argues that ‘merit pay works’ or ‘merit pay doesn’t work’ hasn’t looked closely enough at the evidence. Sensible policymakers should now begin to recognize that some merit pay plans work, and others do not. The challenge is to accumulate sufficient evidence to enable us to tell the difference at the outset.

In terms of teacher attitudes, there is much more to be learned. We know that on average, the kinds of test-based merit-pay schemes generally favored by economists receive little support among the teaching profession. However, there is enough heterogeneity in teacher attitudes that opt-in merit-pay schemes may be feasible. It would also be useful to know more about the kinds of teachers who support merit pay.

As a professor-turned-policymaker, my future academic forays are increasingly likely to be confined to analyzing the work of others. But as a keen follower of the merit-pay literature, I have strong views about the questions I would like to see it answer. Therefore, I conclude with 10 questions that I (and perhaps other policymakers) would like future researchers to answer. These fall into three categories: the impact of merit pay, teacher attitudes towards merit pay, and the politics of merit pay.

### **Impact**

- (1) Which channel is most important to the effectiveness of merit pay: incentive effects or selection effects?
- (2) Are simple merit-pay plans more effective than complex merit-pay plans?<sup>11</sup>
- (3) What is the optimal mix of incentive size and probability of bonus receipt?
- (4) Are teacher-level incentives more effective than school-level incentives?
- (5) Are opt-in plans as effective as compulsory plans?

---

<sup>11</sup> In analysing the impact of simple merit-pay plans, there may be value in researchers studying the impact of the merit-pay schemes that operated in England and Australia in the 19th century.

### Teacher attitudes

- (6) How well do teachers understand their merit-pay plan? When faced with a complex plan, what rules of thumb do they adopt?
- (7) Are teachers who are more effective in the classroom more likely to support merit pay?
- (8) Among potential teachers, is merit pay most attractive to those who would make the best teachers?

### Politics

- (9) What factors predict the longevity of a merit-pay plan? Is there a Nixon-to-China effect, with political parties of the left more likely to implement sustainable merit-pay plans?
- (10) What does the adoption of merit pay mean for other areas of schools reform, such as hiring and firing policies, and the amount of flexibility in the curriculum?

### Acknowledgements

I am grateful to editor Panu Poutvaara, an anonymous referee, Ben Jensen, Gabrielle Leigh, Phillip Metaxas and participants in the 2011 Economics of Education Conference at CESifo, Munich and a seminar at the Grattan Institute for valuable comments on an earlier draft. Particular thanks to Eric Hanushek and Ludger Woessmann for inviting me to prepare this paper for CESifo, thereby providing the unusual opportunity of drawing some links between my previous and current occupations.

### References

- Adkins, G. K. (2004), “Teacher Performance Pay: The Perceptions of Certified School-Based Personnel”, *Doctor of Education Thesis*, College of Education, University of Central Florida, Orlando, FL.
- Atkinson, A., S. Burgess, B. Croxson, P. Gregg, C. Propper, H. Slater and D. Wilson (2009), “Evaluating the Impact of Performance-related pay for Teachers in England”, *Labour Economics* **16**, 251–261.
- Ballou, D. and M. Podgursky (1993), “Teachers’ Attitudes Toward Merit Pay: Examining Conventional Wisdom”, *Industrial and Labor Relations Review* **47**, 50–61.
- Bewley, T. (2004), “Fairness, Reciprocity, and Wage Rigidity”, *IZA Discussion Paper 1137*, IZA, Bonn.

- Buck, S. and J. P. Greene (2010), “Blocking, Diluting, and Co-Opting Merit Pay”, *PEPG Working Paper 10-14*, Harvard University, Cambridge, MA.
- Bushaw, W. J. and S. J. Lopez (2010), “A Time for Change. The 42nd Annual Phi Delta Kappa/Gallup Poll of the Public’s Attitudes Toward the Public Schools”, *Kappan Magazine* **92**, 9–26.
- Campbell, D. T. (1976), *Assessing the Impact of Planned Social Change*, The Public Affairs Center, Dartmouth College, Hanover, NH.
- Chevalier, A., P. Dolton and S. McIntosh (2003), *Teacher Pay and Performance*, Bedford Way papers, Institute of Education, University of London, London.
- Chingos, M. and M. West (2010), “Do More Effective Teachers Earn More Outside of the Classroom?”, *PEPG Working Paper 10-02*, Harvard University, Cambridge, MA.
- Clotfelter, C. T. and H. F. Ladd (1996), “Recognizing and Rewarding Success in Public Schools”, in H. F. Ladd, ed. *Holding Schools Accountable: Performance Based Reform in Education*, The Brookings Institution, Washington, DC, pp. 23–63.
- Corcoran, S. P., W. N. Evans and R. M. Schwab (2004a), “Changing Labor Market Opportunities for Women and the Quality of Teachers, 1957-2000”, *American Economic Review* **94**, 230–235.
- Corcoran, S. P., W. N. Evans and R. M. Schwab (2004b), “Women, the Labor Market, and the Declining Relative Quality of Teachers”, *Journal of Policy Analysis and Management* **23**, 449–470.
- Crabtree, S. (2004), “Can Merit Pay Raise Teacher Pay Without Raising Eyebrows?”, <http://www.gallup.com/poll/13678/can-merit-plans-raise-teacher-pay-without-raising-eyebrows.aspx> (last accessed 31 August 2011).
- Dee, T. S. and B. J. Keys (2004), “Does Merit Pay Reward Good Teachers? Evidence From a Randomized Experiment”, *Journal of Policy Analysis and Management* **23**, 471–488.
- Duffett, A., S. Farkas, A. J. Rotherham and E. Silva (2008), *Waiting to Be Won Over: Teachers Speak on the Profession, Unions, and Reform*, Education Sector, Washington, DC.
- Eberts, R., K. Hollenbeck and J. Stone (2002), “Teacher Performance Incentives and Student Outcomes”, *Journal of Human Resources* **37**, 913–927.
- Elam, S. M. (1989), “The Second Gallup/Phi Delta Kappa Poll of Teachers’ Attitudes Toward the Public Schools”, *Phi Delta Kappan* **70**, 785–798.

- Farkas, S., J. Johnson, A. Duffett, L. Moye and J. Vine (2003), *Stand by Me: What Teachers Really Think about Unions, Merit Pay and Other Professional Matters*, Public Agenda, New York, NY.
- Figlio, D. N. and L. W. Kenny (2007), “Individual Teacher Incentives and Student Performance”, *Journal of Public Economics* **91**, 901–914.
- Fryer, R. G. (2011), “Teacher Incentives and Student Achievement: Evidence from New York City Public Schools”, *NBER Working Paper No. 16850*, NBER, Cambridge, MA.
- Gallup, G. H. (1983), “The 15th Annual Gallup Poll of the Public’s Attitudes Toward the Public Schools”, *Phi Delta Kappan* **65**, 33–47.
- Glewwe, P., N. Ilias and M. Kremer (2010), “Teacher Incentives”, *American Economic Journal: Applied Economics* **2**, 205–227.
- Goldhaber, D., M. DeArmond and S. DeBurgomaster (2011), “Teacher Attitudes About Compensation Reform: Implications For Reform Implementation”, *Industrial and Labor Relations Review* **64**, 441–463.
- Hanushek, E. A. (1998), “The Evidence on Class Size”, *W. Allen Wallis Institute of Political Economy Occasional Paper Number 98-1*, University of Rochester, Rochester, NY.
- Hanushek, E. A. (2011), “The Economic Value of Higher Teacher Quality”, *Economics of Education Review* **30**, 466–479.
- Hanushek, E. A. and S. G. Rivkin (2004), “How to Improve the Supply of High Quality Teachers”, in D. Ravitch, ed. *Brookings Papers on Education Policy*, Brookings Institution Press, Washington, DC, pp. 7–25.
- Hanushek, E. A. and S. G. Rivkin (2006), “Teacher Quality”, in E. A. Hanushek and F. Welch, eds., *Handbook of the Economics of Education, Vol. 2*, North Holland, Amsterdam, pp. 1051–1078.
- Hatry, H. P., J. M. Greiner and B. G. Ashford (1994), *Issues and Case Studies in Teacher Incentive Plans*, 2nd edition, Urban Institute Press, Washington, DC.
- Howell, W., M. West and P. Peterson (2007), “What Americans Think about their Schools”, *Education Next* **7**, 12–26.
- Howell, W., M. West and P. Peterson (2008), “What Americans Think about their Schools, Results from the Second Annual Education Next-PEPG Survey”, *Education Next* **8**, 12–27.
- Howell, W., M. West and P. Peterson (2009), “The Persuadable Public, Results from the Third Annual Education Next-PEPG Survey”, *Education Next* **9**, 20–29.

- Howell, W., M. West and P. Peterson (2010), “Meeting of the Minds, Results from the Fourth Annual Education Next-PEPG Survey”, *Education Next* **11**, 20–31.
- Howell, W., M. West and P. Peterson (2011), “The Public Weighs In on School Reform”, *Education Next* **11**, 11–22.
- Hoxby, C. M. (2000), “The Effects of Class Size on Student Achievement: New Evidence from Population Variation”, *Quarterly Journal of Economics* **115**, 1239–1285.
- Hoxby, C. M. (2002), “The Cost of Accountability”, in W. Evers and H. Walberg, eds., *School Accountability*, Hoover Press, Stanford, pp. 47–73.
- Jacob, B. and L. Lefgren (2008), “Can Principals Identify Effective Teachers? Evidence on Subjective Performance Evaluation in Education”, *Journal of Labor Economics* **26**, 101–136.
- Jacob, B. and M. Springer (2007), “Teacher Attitudes on Pay for Performance: A Pilot Study”, *Working Paper 2007-06*, National Center on Performance Incentives, Peabody College of Vanderbilt University, Nashville, TN.
- Jensen, B. and J. Reichl (2011), *Better Teacher Appraisal and Feedback: Improving Performance*, Grattan Institute, Melbourne.
- Krueger, A. (1999), “Experimental Evidence of Education Production Functions”, *Quarterly Journal of Economics* **114**, 497–532.
- Ladd, H. F. (1999), “The Dallas School Accountability and Incentive Program: An Evaluation of Its Impacts on Student Outcomes”, *Economics of Education Review* **18**, 1–16.
- Lavy, V. (2002), “Evaluating the Effect of Teachers’ Group Performance Incentives on Pupil Achievement”, *Journal of Political Economy* **110**, 1286–1317.
- Lavy, V. (2009), “Performance Pay and Teachers’ Effort, Productivity, and Grading Ethics”, *American Economic Review* **99**, 1979–2011.
- Lazear, E. (2003), “Teacher Incentives”, *Swedish Economic Policy Review* **10**, 197–213.
- Leger Marketing (2011), “Quebecers’ Perceptions Toward Teacher Remuneration - August 2011”, Project: 13026-023, Leger Marketing, Toronto.
- Leigh, A. (2010), “Estimating Teacher Effectiveness From Two-Year Changes in Students’ Test Scores”, *Economics of Education Review* **29**, 480–488.

- Leigh, A. and C. Ryan (2008), “How and Why has Teacher Quality Changed in Australia?”, *Australian Economic Review* **41**, 141–159.
- Marsden, D. (2000), “Teachers Before the Threshold”, Centre for Economic Performance Discussion Paper 454, Centre for Economic Performance, London.
- Marsden, D. and R. Belfield (2006), “Pay for Performance Where Output is Hard to Measure: The Case of Performance Pay for School Teachers”, in B. E. Kauffman and D. Lewin, eds., *Advances in Industrial and Labor Relations*, Vol. 15, JAI Press, London, pp. 1–34.
- Marsh, J. A., M. G. Springer, D. F. McCaffrey, K. Yuan, S. Epstein, J. Koppich, N. Kalra, C. DiMartino and A. Peng (2011), “A Big Apple for Educators: New York City’s Experiment with Schoolwide Performance Bonuses”, *Final Evaluation Report*, Fund for Public Schools, RAND Corporation, Santa Monica, CA.
- McKenzie, P., J. Kos, M. Walker and J. Hong (2008), *Staff in Australia’s Schools 2007*, DEEWR, Canberra.
- Muralidharan, K. (2011), “Long-Term Effects of Teacher Performance Pay: Experimental Evidence from India”, *Mimeo*, University of California San Diego.
- Muralidharan, K. and V. Sundararaman (2011a), “Teacher Performance Pay: Experimental Evidence from India”, *Journal of Political Economy* **119**, 39–77.
- Muralidharan, K. and V. Sundararaman (2011b), “Teacher Opinions on Performance Pay: Evidence from India”, *Economics of Education Review* **30**, 394–403.
- Nadler, C. and M. Wiswall (2009), “Risk Aversion and Support for Merit Pay: Theory and Evidence from Minnesota’s Q Comp Program”, *IESP Working Paper 09-05*, New York University, New York, NY.
- Neal, D. (2009), “Designing Incentive Systems for Schools”, in M. G. Springer, ed. *Performance Incentives: Their Growing Impact on American K-12 Education*, Brookings Institution Press, Washington DC, pp. 149–170.
- Perez, M. (2010), “Inequity Aversion Differences: Experimental Evidence Among Prospective Teachers and Lawyers”, *Mimeo*, Stanford University.
- Perez, M., M. Muraki and S. Loeb (2011), “Do School and Teacher Value-Added Measures Correlate with Their Performance-Based Pay Preferences?”, *Mimeo*, Stanford University.



- Podgursky, M. J. and M. G. Springer (2007), “Teacher Performance Pay: A Review”, *Journal of Policy Analysis and Management* **26**, 909–949.
- Rist, M. C. (1983), “Our Nationwide Poll: Most Teachers Endorse the Merit Pay Concept”, *American School Board Journal* **170**, 23–27.
- Rockoff, J. (2009), “Field Experiments in Class Size from the Early Twentieth Century”, *Journal of Economic Perspectives* **23**, 211–230.
- Rose, L. and A. Gallup (2000), “The 32nd Annual Phi Delta Kappa/ Gallup Poll of the Public’s Attitudes Toward the Public Schools”, *Phi Delta Kappan* **82**, 41–58.
- Rothstein, R. (2009), “The Influence of Scholarship and Experience in Other Fields on Teacher Compensation Reform”, in M. G. Springer, ed. *Performance Incentives: Their Growing Impact on American K-12 Education*, Brookings Institution Press, Washington DC, pp. 87–110.
- Slater, H., N. Davies and S. Burgess (2009), “Do Teachers Matter? Measuring the Variation in Teacher Effectiveness in England”, *Working Paper No. 09/212*, Centre for Market and Public Organisation, Bristol Institute of Public Affairs.
- Springer, M. (2009), “Rethinking Teacher Compensation Policies: Why Now, Why Again?”, in M. G. Springer, ed. *Performance Incentives: Their Growing Impact on American K-12 Education*, Brookings Institution Press, Washington DC, pp. 1–22.
- Springer, M. G., D. Ballou, L. S. Hamilton, V.-N. Le, J. R. Lockwood, D. F. McCaffrey, M. Pepper and B. M. Stecher (2010), “Teacher Pay for Performance: Experimental Evidence from the Project on Incentives in Teaching”, *Conference Paper*, National Center on Performance Incentives.
- Staiger, D. O. and J. E. Rockoff (2010), “Searching for Effective Teachers with Imperfect Information”, *Journal of Economic Perspectives* **24**, 97–118.
- Time Magazine (2010), “TIME Poll Results: Americans’ Views on Teacher Tenure, Merit Pay and Other Education Reforms”, *Time Magazine*, 9 September 2010.
- Vigdor, J. L. (2008), “Teacher Salary Bonuses in North Carolina”, *Conference Paper*, National Center on Performance Incentives, Nashville, TN.
- Winters, M. A., G. W. Ritter, J. H. Barnett and J. P. Greene (2007), “An Evaluation of Teacher Performance Pay in Arkansas”, Heartland Institute Working Paper.



Witham, P. J., J. L. Lewis, M. G. Springer and J. W. Guthrie (2008), *Evaluation of Butler County's PayPLUS Salary Augmentation System*, Management Analysis and Planning Inc, Davis, CA.

Woessmann, L. (2011), "Cross-country Evidence on Teacher Performance Pay", *Economics of Education Review* **30**, 404–418.