

TEACHER COMPENSATION SYSTEMS IN THE UNITED STATES K-12 PUBLIC SCHOOL SYSTEM

Michael Podgursky and Matthew Springer

This paper provides a review of the current teacher compensation system and examines the structure of teacher compensation in the U.S. K-12 public education system. Teacher salaries are largely set by schedules that are neither performance-related nor market-driven, and have significant consequences on school staffing and workforce quality. The second section summarizes the recent literature on compensation reform, with an emphasis on studies using experimental or quasi-experimental designs to evaluate the impact of programs on student achievement and teacher outcomes. A final section offers observations on prospects for future research and reforms.

Keywords: teacher salaries, incentive systems, merit pay, teacher incentives, student performance

JEL Codes: H11, I2, J33, M52

I. INTRODUCTION

During the 2006–07 school year, the most recent year for which national data are available, U.S. public schools spent \$197 billion on salaries and \$64 billion on benefits for instructional personnel (U.S. Department of Education, 2009). These compensation payments account for 55 percent of current expenditures in K-12 public schools and 90 percent of instructional expenditures. As large as these expenditures are, they do not fully capture the resources committed to K-12 compensation, as they do not include compensation for non-instruction personnel or the unfunded liabilities of pension funds and retiree health insurance for teachers and administrators, which recent estimates project to be between \$332 billion and \$933 billion (Pew Center on the States, 2007; Clark, 2009; Barrow and Buck, 2010; Pew Center on the States, 2010).

Michael Podgursky: Department of Economics, University of Missouri, Columbia, MI, USA and Fellow, George W. Bush Institute, Dallas, TX, USA (podgurskym@missouri.edu)

Matthew Springer: Peabody College of Education and Human Development, Vanderbilt University, Nashville, TN, USA and Fellow, George W. Bush Institute, Dallas, TX, USA (matthew.g.springer@vanderbilt.edu)

Teacher salaries and benefits are critical to the total costs and productivity of the K-12 public school system. For instance, if productivity doubles for an input accounting for one percent of total cost, there will be little overall efficiency gain. However, given the large share of costs that arise from teacher compensation, even modest gains in efficiency mean substantial benefits for students, taxpayers, and other education stakeholders.

Teacher compensation is the sum of four parts — base pay, supplements, benefits, and deferred compensation. Base pay is commonly set by salary schedules that have evolved from generations of collective bargaining agreements, or in non-bargaining states like Texas, legislative fiat. Base pay is often augmented by various types of district or state-wide salary supplements (e.g., for coaching an athletic team, mentoring novice teachers, or participating in a career ladder program).¹ Along with fringe benefits such as health insurance and paid leave, deferred compensation also comes in the form of retirement pay.²

An efficient teacher compensation structure is one that is designed to recruit, retain, and motivate the highest quality workforce for any given level of expenditure. However, the current teacher compensation “system” is best characterized as a mix of policies reflecting divergent stakeholder preferences, legislative tinkering, and legacies from earlier vintages of employment contracts. In an effort to promote strategic reform of these systems, policymakers in Washington DC, as well as states and local school districts, have turned their attention to performance-related and market-driven pay plans.

In this paper we provide a critical review of current district pay practices and then offer a descriptive summary of national levels and trends of incentives from various administrations of the U.S. Department of Education’s Schools and Staffing Survey (SASS) (see Table 2 notes). We survey evidence accumulated to date regarding the effectiveness of recent international and domestic policy initiatives to reform teacher compensation systems. The final section discusses future directions for teacher compensation policy reforms and research.

II. THE SINGLE SALARY SCHEDULE AND ITS CONSEQUENCES

In current practice, the most important determinant of a teacher’s pay is the salary schedule in the school district. District salary schedules have been nearly universal in the public school system since the early 1950s, though some locations (primarily southern states) have state-wide teacher salary schedules that establish a minimum pay level but allow for local districts to supplement these minimums. During the 2003–04 school year,

¹ In some states, a significant amount of public funds continue to be allocated for teachers that were grandfathered into now defunct compensation structures. For example, even though the Tennessee legislature repealed its career ladder program in 1997 because nearly all teachers who attempted to qualify for an incentive payment did so, nearly \$70 million was earmarked for the estimated 30,000 career ladder educators still employed by the public school system in the 2008–09 school year.

² The fiscal effects of teacher pension systems are coming under scrutiny as well (Costrell and Podgursky, 2009).

approximately 96 percent of public school districts accounting for nearly 100 percent of all public school teachers reported use of a salary schedule (Podgursky, 2009).

Table 1 displays the 2007–08 salary schedule for teachers in Metropolitan Nashville Public Schools (MNPS). Rows represent years of teaching experience, which range in this case from 0 to 25. Five blocked columns identify post-secondary degrees (e.g., Bachelor, Master, Master +, EDS, and PhD). Similar to most teacher compensation systems in U.S. public school districts, the salary schedule in MNPS provides larger salaries to teachers with higher levels of formal education and for each additional year of teaching experience.

Single salary schedules for teachers contrast with pay practices in most other professions where merit or performance-related pay is more commonplace. In medicine, for

Table 1
Salary Schedule for Metropolitan Nashville Public School District,
2007–08 School Year

Years Experience	Bachelor	Master	Master +	EDS	PhD	Years Experience
0	\$34,059	\$37,665	\$41,512	\$42,932	\$44,718	0
1	34,620	37,986	41,833	43,243	45,038	1
2	35,261	39,108	42,954	44,364	46,160	2
3	36,543	40,390	43,596	45,046	47,442	3
4	37,184	41,031	44,878	46,373	48,725	4
5	38,467	42,313	46,160	47,660	50,007	5
6	39,749	43,596	47,442	49,002	51,289	6
7	41,031	44,878	48,725	50,335	52,571	7
8	42,313	46,160	50,007	51,682	53,853	8
9	43,596	47,442	51,289	53,009	55,136	9
10	44,878	48,725	52,571	54,276	56,418	10
11	46,160	50,007	53,853	55,588	57,700	11
12	47,442	51,289	55,136	56,896	58,982	12
13	48,725	52,571	56,418	58,198	60,265	13
14	50,007	53,853	57,700	59,490	61,547	14
15	50,488	55,136	58,982	60,827	62,829	15
16		55,216	60,265	62,110	64,111	16
17			60,345	62,230	65,394	17
18				62,230		18
19				62,270		19
25	51,369	56,498	61,547	63,472		

Source: Nashville Public School System (www.hr.mnps.org).

instance, pay of doctors and nurses varies by specialty. Even within the same hospital or HMO, pay differs by specialty field (Folland, Goodman, and Stano, 2006). Similarly, in higher education, large differences exist in pay between faculty by teaching field (Ehrenberg, 2004; Oklahoma State University, 2010). Faculty pay structures also tend to be flexible. Starting pay is generally market-driven as institutions often match counter-offers for the more senior faculty they wish to retain. Studies report generally similar findings for private K–12 education and for public charter schools (Ballou and Podgursky, 1997; Ballou, 2001; Podgursky, 2007). Even when private schools report the use of a salary schedule to determine teacher pay levels, payments “off schedule” are frequent. And, unlike the public K-12 system, collective bargaining agreements in higher education often include provisions that allow for field differentials based on external labor market conditions (Rhoades, 1998). Ultimately, the flexibility typically found in pay practices in other fields allow for greater overall cost effectiveness.³

Salary schedules would not be as costly if the factors rewarded, teacher experience and graduate education, were strong predictors of teacher productivity. However, surveys of the education production function literature find little support for a non-subject specific master’s degree positively impacting student achievement, and teacher experience has little effect beyond the first few years (Rivkin, Hanushek, and Kain, 2005; Clotfelter, Ladd, and Vigdor, 2006; Hanushek et al., 2005; Aaronson, Barrow, and Sanders, 2007). Hanushek (2003) reports that out of 41 studies of the effect of a teacher’s education level on their teaching effectiveness (primarily master’s degrees) not a single one found a statistically significant positive effect. In fact, 10 of the studies found statistically significant negative relationships.⁴

There is an old adage in economics: “You can’t repeal the law of supply and demand.” By this economists mean that if governments or regulatory agencies do not allow prices to clear a market then some other mechanism will. School district salary schedules are a case in point. Salaries set by the schedules take no recognition of market or performance factors. Thus, non-price factors act to clear the market. We briefly consider three consequences of these rigid schedules: teacher shortages by field, the concentration of novice teachers in high-poverty schools, and the incentives (or lack thereof) for more effective teachers to stay in classrooms or enter the profession.

A. Shortages by Field

The training, working conditions, and non-teaching opportunities for teachers differ significantly by teaching field, yet the salary schedule within a school district treats all teachers the same, regardless of field. On average the non-teaching opportunities for a

³ For an early critique of teacher salary schedules, with historical background, see Kershaw and McKean (1962).

⁴ For state-by-state estimates of the cost of MA degrees, see Roza and Miller (2009).

high school physical science teacher (or teachers in other technical or scientific fields) are likely more remunerative than for elementary education teachers, yet the salary schedule within a school district gives them identical salaries. Since the salary schedule is rigid, the market clears on quality.

Data from the school principal form of the nationally representative SASS conducted by the U.S. Department of Education illustrates the consequences of these rigidities. Principal respondents were asked a series of questions about how difficult or easy it was for them to fill teaching vacancies by field, based on a four point scale ranging from “easy” to “could not fill the vacancy.” Approximately 75 percent of respondents who needed to hire an elementary education teacher reported it was “easy” to fill the vacancy in the 2003–04 school year. In contrast, only 30 to 35 percent of principal respondents who needed to fill a science, mathematics, or special education opening gave such an assessment.

Survey data from schools shows that science, mathematics, and special education teachers tend to be less likely to have majored in their primary field of instruction and are more likely to be classified as teaching “out of field” than elementary school teachers (Podgursky, 2009, 2010; Seastrom, Gruber, Henke, McGrath, and Cohen, 2004). For example, Ingersoll (1999) estimates that one-third of all secondary school teachers who teach mathematics do not have either a major or minor in math or a related discipline, while more than one-half of teachers tasked with leading a physical science class do not have a major or minor in the field. While there are numerous forces at work in the labor market for teachers, these general patterns align with empirical evidence that the quality of teachers has declined relative to that of highly skilled workers, college-educated workers, and the overall labor force (Bacolod, 2001; Corcoran, Evans, and Schwab, 2004; Grogger and Eide, 1995).

B. Inequitable Distribution of Teacher Quality within Districts

The inequitable distribution of high quality teachers across schools helps to explain the student achievement gap reported by many urban school systems. Most public schools differ in attractiveness as places to teach, with schools that have higher concentrations of low-income, non-white, and low-performing students being perceived as less desirable places to work. More experienced teachers can typically use seniority-based transfer provisions in collective bargaining agreements to choose where to teach, and they can be expected to use it to exit these less desirable placements.⁵ Transfer rights contribute to the disparities in quality teachers across schools because restrictive contracts put low-income, non-white, and low-performing schools at a disadvantage in the competition for teachers and resources within districts (Moe, 2009).

⁵ In a case study of five large, urban school systems, Levin, Mulhern, and Schunck (2005) reported that 40 percent of teaching vacancies were filled by incumbent teachers and that school administrators had very little or no voice in the hiring decision.

This sorting of teachers across schools further strengthens racial and poverty-related achievement gaps. Schools enrolling children from the most disadvantaged backgrounds are more likely to be staffed by teachers graduating from less competitive colleges, teachers instructing out-of-field, and novice teachers (Lankford, Loeb, and Wyckoff, 2002; Iaterola and Steifel, 2003; Roza et al., 2007). Teacher effectiveness research consistently finds that novice teachers (i.e., first or second year teachers) produce smaller achievement gains for their students than more experienced teachers (Aaronson, Barrow, and Sander, 2007; Rivkin, Hanushek, and Kane, 2005). The net result is that children enrolled in schools with high concentrations of disadvantaged students have greater exposure to less qualified instructors.

The inequitable distribution of high-quality teachers among schools within districts is arguably a consequence of uniform teacher salary schedules in conjunction with differences in nonpecuniary characteristics of schools (e.g., condition of school building, principal leadership, safety, and distance from home). When pay is equalized, teacher quality is disequalized across schools. In order to equalize teacher quality, schools have begun to experiment with incentives designed to help offset differences in non-wage job characteristics across schools (Prince, 2002, 2003; Kirshtein et al., 2004; Steele, Murnane, and Willett, 2010). Unfortunately, there is little research on the compensating differential needed to offset differences in nonpecuniary workplace characteristics. It is also likely that the required bonuses ("combat pay") would vary from district to district depending on relative school characteristics and teacher preferences.

C. Lack of Incentives for More Effective Teachers to Stay on the Job or Enter the Profession

Some teachers are consistently better at raising the achievement of their students than others. Value-added studies of teacher effectiveness consistently find large variation in teacher classroom performance (Aaronson, Barrow, and Sander, 2007; Hanushek et al., 2005; Kane, Rockoff, and Staiger, 2008; Sanders and Rivers, 1996; Rockoff, 2004). Top-performing teachers, as defined by those teachers at the 95th percentile, produce three times the achievement growth in students when compared to low-performing teachers (Hanushek, 2003). Hanushek and Rivkin (2004) reported that the achievement gap among high- and low-socioeconomic status students could be overcome if an economically disadvantaged student encountered an above average teacher for five consecutive years.

In a related study, Goldhaber and Hansen (2010) analyze the effect of using teacher value-added estimates to guide teacher tenure decisions. They report that if teachers with early career value-added estimates in the bottom 25 percent of the distribution are dismissed, there is an educationally significant effect on the distribution of teacher quality. Whether or not performance pay raises teacher effectiveness (a point taken up below), if a compensation scheme could induce highly effective teachers to stay

and ineffective teachers to leave, workforce quality and student achievement would improve.⁶

III. TRENDS IN COMPENSATION REFORM

Despite the fact that compensation payments account for a majority of the total budget of a school system, relatively little microeconomic data on the design of these programs or information about the experiences of school systems implementing various pay reform models exists. The best data currently available on national levels and trends comes from the SASS, which samples schools, principals, and teachers using a stratified probability design. The SASS has been fielded every four to five years starting in the 1987–88 school year.^{7,8} However, due to the highly skewed size distribution of school districts we report most statistics based on the SASS data in two ways: as a percentage of all school districts and as a percentage of all teachers.

Table 2 displays estimates from a series of items that asked a school district official if their district provided pay bonuses or other rewards for certain teacher characteristics or behaviors.⁹ As of the 2007–08 school year, 24.5 percent of administrators offered a bonus to teachers with National Board for Professional Teaching Standards (NBPTS) certification. Administrators offering these bonuses were clustered in districts well above average in size, as can be deduced from the teacher-weighted estimate being nearly two times greater (e.g., 48.5 percent of all public school teachers in the U.S. were exposed to a program that offered a bonus if they earned NBPTS certification). NBPTS certification is the most rapidly growing form of incentive pay reported by respondents, with exposure rising by 26 percentage points from the 1999–2000 to 2007–08 school years.

Incentive payments for excellence in teaching, teaching in a less desirable location, or teaching in a shortage field are less popular, although the incidence of these district

⁶ For a discussion of how much progress in student achievement could be accomplished by instituting a program of removing the least effective teachers, see Hanushek (2009).

⁷ SASS includes private schools and teachers as well. However, the focus of this study is on trends in public schools. Though the SASS covers two decades of public school experience and has included various questions about market-driven and performance-related pay, many of the compensation-specific survey questions are longitudinally inconsistent. Thus, we focus attention on data in the most recent waves of the survey, which have maintained some consistency.

⁸ Several researchers have relied on the SASS data to study educator compensation, incentives, and the labor market. Ballou (2001) uses the SASS data to compare bonus pay in public versus private schools. His analysis of teacher salary data suggests that the size of bonuses in public schools is small, both as a share of salary and in comparison to private schools. In addition to Ballou, some other researchers have analyzed factors associated with the use of performance pay in public schools. Goldhaber et al. (2005) examine factors associated with the incidence of performance pay in public schools. Podgursky (2009) shows charter schools are much more likely to make use of market-driven or performance-related pay incentives.

⁹ The wording of the survey question was as follows, “Does the district currently use any pay incentives such as a cash bonuses, salary increase, or different steps on a salary schedule to reward ...?”

Table 2
Incidence of Various Incentive Payment Schemes

District Rewards Following:	District-Weighted (%)				Teacher-Weighted (%)			
	1999-2000	2003-04	2007-08	Change	1999-2000	2003-04	2007-08	Change
NBPTS	8.3	18.4	24.5	16.2	22.9	39.8	48.5	25.6
Excellence in teaching	5.5	8.0	10.3	4.8	13.6	14.0	15.3	1.7
Teach in less desirable location	3.6	4.6	5.7	2.1	11.2	13.1	16.1	4.9
Teach in fields of shortage	10.4	11.9	15.4	5.0	23.6	25.3	32.2	8.6
In-service professional development	26.4	24.2			38.8	35.9		
Number of incentives								
(excluding in-serv. prof. dev.)								
None	78.0	68.7	61.0	-17.0	56.46	41.2	36.1	-20.4
1 Incentive	17.1	23.4	27.0	9.9	26.92	36.3	35.3	8.4
2 Incentives	4.0	6.2	8.0	4.0	9.8	14.4	14.5	4.7
3 Incentives	0.7	1.5	3.1	2.4	2.52	6.8	8.8	6.2
4 Incentives	0.2	0.3	0.9	0.7	4.3	1.3	5.4	1.1
Based on student achievement, were any schools in the district rewarded in any of the following ways?								
Cash bonus/add resources for school-wide activity			6.8				19.6	
Cash bonus/add resources for teachers			4.7				15.4	
Schools given non-monetary forms of recognition			15.8				30.4	

Source: Calculated from U.S. Department of Education's *Schools and Staffing Surveys*, various years, <http://nces.ed.gov/surveys/sass/>.

rewards has increased modestly since the 1999–2000 school year. Nearly one-third of all public school teachers were employed in a school district that offered some type of incentive for teaching in a field with shortages during the 2007–08 school year, an increase of 8.6 percentage points from data collected during 1999–2000 school year.

The middle rows of Table 2 display estimates of the number of district rewards provided to teachers, excluding in-service professional development. Only five percent of teachers were employed in public school districts where the respondent reported all four incentives being present during the 2007–08 school year. Further, 61 percent of school districts, employing an estimated 36.1 percent of all K–12 public school teachers in the United States, did not offer any incentives for earning NBPTS certification, excellence in teaching, teaching in a less desirable location, or teaching in fields of shortage, although the share of districts with none of these types of incentives dropped by 17 percentage points since the 1999–2000 school year.

The bottom rows of Table 2 display summary statistics for three questions about team level reward programs. Of most interest in the context of this study is the item regarding teachers being awarded cash bonuses and/or additional resources if their school was recognized based on student achievement. Approximately five percent of public school districts offered this type of incentive payment during the 2003–04 school year, accounting for 15.4 percent of all public school teachers. Unfortunately, we cannot examine whether the incidence of these group incentive programs changed over time since this battery of questions only appeared on the 2003–04 school principal form of the SASS.

The two most recent waves of the SASS asked school district administrators about various methods used in their district to recruit teachers. As displayed in Table 3, loan forgiveness programs were the most prevalent strategy used by districts to recruit teachers during the 2003–04 school year. However, we cannot study the trend over time because the questions were inconsistent between administrations (i.e., the 2007–08 form asks about “forgiveness of student loans funded by the district” as opposed to “student loan forgiveness” more generally).

Roughly seven percent of school districts employing 15.9 percent of all public school teachers used some form of signing bonus to assist with the recruitment of teachers during the 2007–08 school year. Even though the number of school districts offering signing bonuses increased by 2.1 percentage points over time, there was a very slight decrease in the percentage of teachers employed in those districts. These subtle shifts align with the growing number of rural districts exploring financial incentives to fill vacancies.¹⁰

While all of the SASS surveys included questions about performance-related and market-driven pay, only a few of the questions were consistently asked from one survey administration to the next. One block of questions that was nearly identical concerned teacher recruitment bonuses by field. School district administrators were asked whether

¹⁰ Table 3 also indicates that slightly more than three percent of school districts provided relocation assistance to new hires. Districts offering finder’s fees to existing staff for new teacher referrals were reported in less than two percent of school districts.

Table 3
District Use of Various Recruitment Incentives

Methods district uses to recruit teachers:	District-Weighted (%)			Teacher-Weighted (%)		
	2003–04	2007–08	Change	2003–04	2007–08	Change
Signing bonuses	4.8	6.8	2.1	17.4	15.9	–1.5
Student loan forgiveness	8.1			19.0		
Forgiveness of student loan(s) funded by district		2.3			5.1	
Relocation assistance	3.2	3.6	0.4	8.5	9.2	0.7
Finder's fee to existing staff for new teacher referrals	0.9	1.6	0.6	2.3	2.5	0.2
Available training to staff members to teach for current or anticipated shortage		30.7			38.3	

Source: Calculated from U.S. Department of Education's *Schools and Staffing Surveys*, various years, <http://nces.ed.gov/surveys/sass/>.

their district currently offered pay incentives to recruit or retain teachers to teach in shortage fields and, if so, to identify the fields in which incentive pay was used. Unfortunately, this set of questions was not included in the district survey administered during the 2007–08 school year.

Table 4 displays summary statistics on the incidence of rewards to recruit and/or retain teachers in fields of shortage. As indicated in the table, there is a sharp increase in the usage of rewards to recruit and/or retain teachers in fields of shortage over the 16 year interval. In the 1987–88 school year, only 7.5 percent of districts, or 11.3 percent of all public school teachers worked for a system that provided such incentives.¹¹ That share climbed to 12 percent of districts employing 25 percent of teachers by the 2003–04 school year. And, consistent with the recruitment difficulty responses displayed in Table 3, these incentives were most commonly reported for special education, mathematics, science, and English as a second language.

Data from the SASS suggest that the national level of performance-related and market-driven pay reform is increasing. However, one serious limitation is that respondents are

¹¹ Note that these recruitment incentives can take the form of cash bonuses, higher pay, or higher initial placement on the salary schedule. The latter is more subtle, and thus less controversial, than explicit bonuses or differentiated pay structures.

Table 4
Incidence of Rewards to Recruit/Retain Teachers in Field of Shortage

	Reward to Recruit/Retain Teachers in Fields of Shortage %					Change
	1987-88	1990-91	1993-94	1999-2000	2003-04	
District provides incentive	11.3	16.6	18.7	23.6	25.3	14.0
Elementary				2.4	2.6	
Special Education	6.7	11.8	13.4	14.3	20.6	13.9
English/Language Arts				5.3	4.2	
Social Studies				1.6	2.4	
Computer Science	1.4	2.9	1.3	3.4	3.4	2.0
Mathematics	5.2	5.8	3.9	8.9	15.7	10.5
Physical Sciences	3.6	5.0	3.9	8.4	13.4	9.8
Biological Sciences	3.8	4.3	3.7	8.4	12.8	8.9
English as Second Language	3.3	7.6	8.1	11.1	15.5	12.2
Foreign Language	2.4	3.1	2.4	5.3	9.4	7.0
Music or Art				4.9	6.4	
Vocational/Technical Education		4.7	3.2	8.0	7.3	
Other Fields	4.2	4.2	1.6			

Source: Calculated from U.S. Department of Education's *Schools and Staffing Surveys*, various years, <http://nces.ed.gov/surveys/sass/>.

never asked about the size of bonuses, either in dollars or as a percent of salary. Even though the education-specific research base is not sufficiently robust to prescribe an optimal incentive award amount, not knowing the size of bonuses is disconcerting. A number of studies have found that monetary awards given to public school teachers are quite small relative to private schools (Ballou and Podgursky, 1997; Figlio and Kenny, 2007).¹² Additionally, several studies report that the amount of the bonus award matters (Figlio and Kenny, 2007; Clotfelter et al., 2008; Taylor and Springer, 2010; Springer et al., 2010a).

IV. EVALUATIONS OF PERFORMANCE-RELATED AND MARKET-DRIVEN PAY REFORMS

This section reviews recent evaluation studies that have assessed the impact of performance-related or market-driven pay programs on student achievement and teacher outcomes. The review focuses on evaluations using randomized designs and rigorous quasi-experimental designs (such as a regression discontinuity framework) because, when implemented properly, they provide the best estimates of the causal effect of an intervention on an educational outcome.^{13,14} We relax this standard in a few cases, for exceptionally well-designed non-experimental studies.

A. International Evidence

Table 5 summarizes key design components of recent teacher compensation reforms that were evaluated using strong experimental designs, as well as the study period, sample size, dependent variable(s), and basic findings. All of these studies were implemented abroad. Though most report generally positive effects on student achievement, it is less clear whether these programs actually promoted long-run learning, as some studies find the effects do not persist over time or document opportunistic behaviors on the part of treatment teachers that account for increased student achievement. Furthermore, the incentive structure facing teachers and schools in several of the studies (e.g., in Andhra Pradesh, India or rural Kenya) are very different from the operational context found within the U.S. public school system. Even so, these studies represent some of the best scientific evidence on this issue.

Muralidharan and Sundararaman (2009) led the design and implementation of the Andhra Pradesh Randomized Evaluation Study (AP REST) in partnership with the gov-

¹² Reviews of the evaluation literature on private sector pay reforms indicate the target payout is between four and 12 percent of base pay (Varadarajan and Futrell, 1984; Lazear, 2000; Prince et. al., 2010).

¹³ Podgursky and Springer (2007) review the earlier literature in this area, comprising largely non-experimental studies using multivariate statistical methods to estimate the treatment (performance-related pay programs) effect on student achievement or similar educational outcomes.

¹⁴ Thistlewaite and Campbell (1960), Hahn, Todd, and van der Klaauw (2001), and Lee and Lemieux (2009) provide discussions of randomized design experiments.

ernment of Andhra Pradesh, the Azim Premji Foundation, and the World Bank. AP RESt randomly selected and assigned 500 rural Indian schools to one of the four treatment conditions or to the control group. The treatment conditions included two output-based incentive systems (an individual teacher incentive program and a group-level teacher incentive program) and two input-based resource interventions (one providing an extra-paraprofessional teacher and another providing block grants).

Muralidharan and Sundararaman report that student test scores on high-stakes tests increased between 0.12 and 0.19 standard deviations in the first year of the program and between 0.16 and 0.27 standard deviations in the second. Students enrolled in classrooms presided over by bonus-eligible teachers scored 0.11 to 0.18 standard deviations higher on low-stakes tests than students whose teachers were not eligible to earn a bonus award. Students in incentive classrooms also scored higher on a separate, “high-order thinking” test, which the authors suggest represents “genuine improvements” in learning, as opposed to better test-taking skills or perhaps other strategies employed by teachers to increase their chances of receiving an AP RESt bonus award.

The schools assigned to the output-based intervention (i.e., individual- or group-incentive conditions) also outperformed those schools assigned to the input-based resource interventions (i.e., paraprofessional or block grant conditions). Additionally, students enrolled in a classroom instructed by a teacher selected for the group incentive intervention also outperformed students in control-condition classrooms on the mathematics and language tests (by 0.28 and 0.16 standard deviations, respectively). Finally, students enrolled in schools assigned to the individual incentive treatment outperformed students in both the group incentive treatment and the control classrooms following the second year of implementation.

Glewwe, Ilias, and Kremer (2008) studied the impact of the International Child Support Incentive Program (ICSIP), a group incentive intervention that randomly assigned 100 schools in rural Kenya to either a treatment or a control condition. Unlike the AP RESt program, ICSIP’s bonus scheme was structured as a rank-ordered tournament and did not offer cash bonuses. Prizes ranging between 21 percent and 43 percent of average monthly base salary were awarded on the basis of student drop-out rates and test scores to the twelve highest-performing and the twelve most-improved treatment group schools.¹⁵

Glewwe, Ilias, and Kremer (2008) find that students enrolled in schools participating in the ICSIP intervention had noticeably higher scores on high-stakes tests than students enrolled in schools assigned to the control condition. However, when comparing the test performance of students enrolled in control and treatment schools on low-stakes exams, there were no systematic differences. It appeared that students enrolled in schools participating in the ICSIP intervention were coached in test-taking skills; an

¹⁵ Unlike other incentive programs discussed in this section of the paper, ICSIP awarded teachers with prizes rather than cash bonuses. As noted by Glewwe, Ilias, and Kremer (2008), the ICSIP awarded prizes such as a suit worth about \$50, plates, glasses and cutlery worth about \$40, a tea set worth about \$30, and bed linens and blankets worth about \$25.

Table 5
Summary of Recent Experimental and Quasi-Experimental Evaluations of International Pay Reforms

Program	Study Design	Study Period	Sample	Performance Measures		
				Unit of Accountability	Performance Measures	Results
Kenya's International Christelijk Steunfonds Incentive Program	RCT	1998-1999	100 primary schools; 1,000+ teachers; 50,842 students.	Group (school)	Student test score gains and student achievement levels	Modest, positive effect for high-stakes assessment (0.14 standard deviations above control group). No effect on low-stakes assessment.
Andra Pradesh, India's Randomized Evaluation Project	RCT	2006-2008	300 schools and 68,000+ student observations.	Individual and Group (school)	Student test score gains	Modest, positive effect on high-stakes assessment (approx. 0.12 to 0.19 standard deviations after year one and 0.16 to 0.19 standard deviations after year two).
Israel's Ministry of Education's School Performance Program	RD	1994-1997	62 schools (37 non-religious, 18 religious and 7 Arab schools).	Group (school)	Number of credit units per student, student receiving a matriculation certification, and school dropout rate.	Modest, positive effect for average credit hours earned (increased 0.70 units), average science credits earned (increased, 0.41 units), average test score (increased 1.75 points), and proportion of students taking Israel's matriculation exam (increased 2.1 percent).
Israeli Teacher-Incentive Experiment	RD	2001	4,109 students and 27 schools.	Individual	Student achievement levels	Modest, positive effect for number of exit exam credits earned in mathematics (increased 18 percent) and in reading (increased 17 percent).

Mexico's Carrera Magisterial	RD	1998–2003	850,000+ classroom-year observations. 810 primary school teachers; 209 secondary school teachers.	Individual	Educational degrees, years of experience, professional development, principal ratings, content knowledge mastery, and student performance on standardized tests.	No effect for primary school teachers; Modest, positive effect for secondary school teachers (approximately 3 to 15 percent of standard deviation).
Mexico's Carrera Magisterial	RD	2000–2002	76,567 teachers and 27,123 schools.	Individual	Educational degrees, years of experience, professional development, principal ratings, content knowledge mastery, and student performance on standardized tests.	Small, positive effects (<10 percent of standard deviation).

Note: RCT denotes randomized controlled trial and RD denotes regression discontinuity.

Sources: Glewwe, Ilias, and Kremer (2008), Lavy (2002), Lavy (2009), McEwan and Santhabáñez (2005), Muralidharan and Sundaraman (2009), and Santhabáñez et al. (2007).

analysis of item-level test data revealed, for example, that students enrolled in treatment group schools were significantly less likely to leave a test question blank. Moreover, the research team also did not detect any systematic differences in teacher attendance or pedagogy (behavior in classroom, instructional practices, number of homework assignments) among teachers in treatment and control schools. Rather, teachers working in schools eligible for an ICSIP prize were 7.4 percentage points more likely to offer test-preparation sessions for students outside of normal school hours (typically when students were on vacation).¹⁶

Lavy (2002) evaluated a group incentive program implemented in 62 Israeli high schools that was designed to reduce student drop-out rates and improve student achievement. The program rewarded school performance on the basis of three factors: mean test scores, mean number of credit hours, and school drop-out rate. The bonus scheme was designed as a rank-ordered tournament, with schools in the top third of performers competing for \$1.44 million in awards. Schools earning a bonus had to distribute 75 percent of the school-level award funds to teachers in amounts proportional to their gross annual compensation, regardless of the teacher's performance during the school year; the remaining 25 percent was to be used for improving school facilities for teachers. In total, top-performing schools received between \$13,000 and \$105,000 during the first year of implementation, with teacher bonuses ranging from \$250 to \$1,000 per teacher.

Lavy reports a positive and statistically significant effect of the program on student outcomes. Following the second year of implementation, for example, the program was found to have a positive effect on average credit hours earned, average science credits earned, average test scores, and the proportion of students taking Israel's matriculation test. Estimates further indicate the program affected particular groups of students more than others—for instance, students at the low end of the ability distribution performed much better than expected on Israel's exit exams.

Lavy (2002) also compared the effectiveness of Israel's group incentive intervention with an input-based intervention that had been implemented several years earlier. The input-based intervention provided 22 secondary schools with additional resources to implement professional training programs, reduce class size, and offer tutoring to below-average students. Although both programs improved student outcomes, Lavy (2002) concludes that the group incentive program is more cost-effective at the margin. Muralidharan and Sundararaman similarly find both the individual and group incentive programs were more cost-effective than either the "extra-paraprofessional" teacher or block-grant treatment conditions. The relative effectiveness of these interventions is particularly relevant to U.S. education policy as input-based reforms generally have been implemented more widely than output-based interventions.¹⁷

¹⁶ Glewwe, Ilias, and Kremer (2008) reported that the exit rate of teachers was not significantly different between program and comparison schools. And, even though the entry rates of teachers into incentive schools were different, they were not statistically significant at conventional levels.

¹⁷ Hanushek (2003) provides a critical review of evidence on input-based schooling policies in the United States and abroad.

Lavy (2009) examined the impact of an individual incentive program in Israel that awarded bonuses to high school teachers in grades ten, eleven, and twelve based on their students' performance on national exit tests. The program was structured as a rank-ordered tournament and operated for a single semester (January–June 2001). Teachers in the intervention could earn a bonus for each class of students they prepared for the national exit tests, with awards ranging from \$1,750 to \$7,500 per class prepared.

In this analysis, Lavy exploited two subtle features of the performance-related pay program — measurement error in the assignment variable and a break along the pre-intervention assignment variable — to estimate the causal impact of the incentive program by using regression discontinuity design. Estimates of the net intervention effect indicated the number of exit-exam credits earned by students instructed by a teacher in the incentive program increased by 18 percent in mathematics and 17 percent in English, while data from a survey of teacher attitudes and behaviors suggested positive changes in teaching practices, teacher effort, and instruction tailored to low-performing students. When investigating gaps in performance between the results of school tests and national tests taken by students enrolled in treatment and comparison schools, Lavy did not find evidence of opportunistic behavior or negative spillover effects.

Santibañez et al. (2007) use a regression discontinuity (RD) design to estimate the impact of Mexico's Carrera Magisterial (CM) on student test scores. Implemented in 1992, CM is a teacher incentive program designed collaboratively by state and federal education departments and the national teachers' union. Teachers participating in the program can earn a financial bonus by accumulating enough points on a variety of CM-defined measures, including input criteria such as years of experience, highest degree held, and professional development activities, as well as output criteria such as their performance on a subject-matter knowledge test and their students' test scores (Santibañez et al., 2007). Awards ranged from 24.5 to 197 percent of a teacher's annual earnings (McEwan and Santibañez, 2005; Ortiz-Jimenez, 2003).

Santibañez et al. (2007) investigate the impact of financial incentives that individual teachers have to improve their students' test performance. Since the program appraises teachers on most performance measures before students take the high-stakes tests each school year, participant teachers have a general sense of how many additional points they need to earn based on their students' performance on the high-stakes test to receive an award. Santibañez et al. (2007) detect a negligible impact on test scores of students enrolled in elementary school classrooms taught by teachers facing a strong incentive, though they detect small, positive effects at the secondary level.

B. Evidence from the United States

Table 6 summarizes key design components of the performance-related pay programs that took place in the U.S. public school system and were evaluated using strong, experimental designs, as well as the study period, sample size, dependent variable(s), and main findings. In August 2006, the National Center on Performance Incentives (NCPI) implemented the Project on Incentives in Teaching (POINT) intervention in the MNPS

Table 6
Summary of Recent Experimental and Quasi-Experimental Evaluations of U.S. Pay Reforms

Program	Study Design	Study Period	Sample	Performance Measures		
				Unit of Accountability	Measures of Teacher Performance	Results
Project on Incentives in Teaching (Nashville, TN)	RCT	2007-2009	147 treatment and 152 control teachers (grades 5 - 8).	Individual	Student test scores in mathematics, reading, social studies, and science.	No overall effect. Modest, positive effect for 5 th graders in years 2 and 3 (approx. 0.5 to 0.66 of a year's typical growth), though effects no longer evident the following school year.
Recognizing Excellence in Academic Leadership Program (Chicago, IL)	RCT	2008-2011	32 Teacher Advancement Program (TAP) schools.	Individual and Group (school)	Mentor review, self-review, master teacher review, administrator review, classroom observations, teacher developed portfolio, interviews, student test score gains, and overall school performance.	No overall effect after two years. Small, positive effect on teacher retention after year 1, though effect no longer evident after 2 years.
School-Wide Performance Bonus Program (New York, NY)	RCT	2008-2009	Approx. 191 treatment and 131 control group schools (elementary, middle and k-8). More than 100,000 in grades 3 through 8.	Group (school)	Student test score levels and gains, student, teacher, and principal perceptions of school environment, and external enumerators rating of school's instructional climate.	No overall effect after two years. Small, marginally significant reduction in number of teacher absences among treatment teachers in small schools (approx. 2.5 fewer absences over 5 month period).
Project on Team-Level Incentives in Teaching (Round Rock, TX)	RCT	2009-2010	39 treatment and 39 control group teams (grades 6 - 8).	Group (grade-level teams)	Student test scores in mathematics, reading, social studies, and science.	In-progress

Note: RCT denotes randomized controlled trial.

Sources: Glazerman, McKie, and Carey (2009), Goodman and Turner (2011), Springer et al. (2010b), and Springer and Winters (2009).

system. The POINT intervention was designed as an individual incentive intervention in which a value-added measure of teacher performance was judged according to a fixed performance standard. Nearly 300 teachers of middle school mathematics volunteered and were randomly assigned to either the treatment or control condition. Teachers assigned to the intervention were eligible to receive bonuses of up to \$15,000 per year for a three-year period on the basis of two factors: the progress of a teacher's math students over a year, as measured by their gains on the Tennessee Comprehensive Assessment Program (TCAP); and the progress of a teacher's non-math students over a year, as measured by their gains on the TCAP as well.

Springer et al. (2010b) report that, while the general trend in middle school mathematics performance was increasing over the period of the project, students of teachers randomly assigned to the treatment group did not outperform students whose teachers were assigned to the control group. Researchers detected a positive effect of incentives in fifth grade during the second and third years of the experiment. This finding, which is robust to a variety of alternative estimation methods, is nonetheless of limited policy significance because this effect does not appear to persist after students leave fifth grade. Students whose fifth grade teacher was in the treatment group performed no better by the end of sixth grade than did sixth graders whose teacher the year before was in the control group.

Glazerman, McKie, and Carey (2009) designed an impact evaluation of the Teacher Advancement Program (TAP). At the beginning of the 2007–08 school year, 16 schools were randomly assigned to either the TAP intervention or the control condition.¹⁸ Another 16 schools were then recruited and randomly assigned to the TAP intervention or control conditions at the beginning of the 2009–10 and 2010–11 school years. After two years of implementation, Glazerman, McKie, and Carey (2009) report an indeterminate effect on both student outcomes as well as teacher behavior.¹⁹

New York City Public Schools implemented the School-Wide Performance Bonus Program (SPBP) midway into the 2007–08 school year. The SPBP was designed to provide financial rewards to teachers in schools serving disadvantaged students. The program sets expected incentive payments as a fixed performance standard, meaning that schools participating in the program are not competing against one another for a fixed sum of money. All participating schools can earn bonus awards of up to \$3,000 per full-time union member working at the school if the school meets predetermined performance targets defined by the New York City Department of Education's accountability program, with the idea that this sum will be used to award bonuses to teachers and

¹⁸ The TAP is a comprehensive school-reform model consisting of four elements: (1) multiple career paths; (2) ongoing, applied professional growth; (3) instructionally focused accountability; and (4) performance-based compensation. More information on the TAP can be found at www.talentedteachers.org. For non-experimental evaluations of the TAP see Springer, Ballou, and Peng (2008) and Hudson (2010).

¹⁹ Quasi-experimental design estimates (nearest-neighbor propensity score methods) indicated that the program had statistically significant positive effect on increasing teacher retention at TAP schools during the first year.

staff found to be deserving. The SPBP rules further mandate that schools participating in the program establish a four-person site-based compensation committee to determine how bonus awards will be distributed to school personnel.

Springer and Winters (2009) examined the impact of the SPBP on student outcomes and the school learning environment using data from the randomized lottery selection process. Their sample included 186 SPBP-eligible elementary, K–8, and middle schools and 137 control-condition schools in New York City. Overall, they find that the SPBP had little impact on student proficiency or school environment after two years of implementation. Goodman and Turner (2011) reach a similar conclusion, while offering some evidence of decreased teacher absenteeism in SPBP schools that with fewer than five teachers in tested grades and subjects. In summer 2011, the RAND Corporation, in partnership with the NCPI, is expected to release findings from a three-year long assessment of the SPBP.

During the 2007–08 school year, the NCPI implemented a demonstration project to evaluate the impact of a team-level performance-related pay program. Of teachers in grades six, seven, or eight in Texas' Round Rock Independent School District, 78 grade-level teams were randomly assigned to either the treatment or control conditions at the beginning of both the 2007–08 and 2008–09 school years. A team was defined as a group of academic teachers who meet regularly to discuss a common set of students, performance goals, and outcomes for which they are collectively accountable. An individual teacher on a team that was assigned to the incentive intervention was eligible for an approximate \$6,000 cash award if their team was one of the four highest-performing teams at their grade level as measured by a team-level value-added score. Results are expected to be released in spring 2011.

With funding from the U.S. Department of Education's Institute of Education Sciences, researchers at Mathematica Policy Research, Inc. designed a randomized experiment to assess the efficacy of the Talent Transfer Initiative (TTI). In seven large, diverse school districts throughout the country, Glazerman, McKie, and Carey (2009) used value added analysis of three years of student achievement growth data to identify the top 20 percent of elementary, middle school English, and middle school math teachers. These high-performing teachers were then offered \$20,000 to transfer to low achieving schools and the research team facilitated their transfers. At the same time, the research team identified a pool of low achieving schools with teaching vacancies in the targeted grades and subjects and half of them were assigned to a treatment group and half to a control group. Treatment schools were eligible to hire a top-tier TTI teacher that had been offered \$20,000 to transfer. Control schools had to fill their vacancies the way they normally would. The study followed the transfer teachers and the corresponding control teachers for two years and the final report is expected to be released in spring 2011.

While Glazerman, McKie, and Carey (2009) are the first research team to directly test the compositional effect of pay incentives in U.S. K–12 public schooling using an experimental design, a number of informative observational studies have recently assessed the effectiveness of similar policies in attracting or retaining teachers. Clotfelter et al. (2008) reported that an \$1,800 retention bonus for mathematics, science, and special education teachers working in low-income or low-performing public schools in North

Carolina reduced turnover among eligible teachers by 17 percent. Taylor and Springer (2010) take advantage of a recent natural experiment in Texas to study the impact of incentive payments on teacher turnover. They find that teachers who received no award had a heightened probability of turnover, while teacher who received relatively large awards had a greatly reduced probability of turnover.

Steele, Murnane, and Willett (2010) studied the impact of the California Governor's Teaching Fellowship, a \$20,000 conditional scholarship designed to attract academically talented, newly licensed teachers to schools in the bottom half of the achievement distribution and to retain them in these low-performing schools for at least four years. They estimate that, in the absence of the program, 28 percent of the fellowship recipients would not have taught in a low performing school. And, approximately 75 percent of those individuals still taught in a low-performing school after four years of service.

In a large-scale evaluation of Texas' District Awards for Teaching Excellence (DATE) program, Springer et al. (2010a) report that school districts with select school incentive pay plans experienced statistically significant declines in the share of teachers who were leaving for other school districts, or leaving teaching altogether. Moreover, teacher turnover was related to the size of the maximum award proposed under district's DATE plans; teacher turnover increased for districts with relatively small proposed maximum awards, and decreased as the proposed maximum award amount increased, until the maximum award exceeded roughly \$6,000.

V. PROSPECTS FOR FUTURE RESEARCH AND REFORMS

Human resource policy — including the recruitment, retention, and motivation of employees — is increasingly recognized as a critical variable in the success of an organization. An integrated and coherent compensation policy is the central core of an efficient human resource policy. In private and many public organizations, the compensation package is considered as a strategic whole, and carefully designed to get the most human resource return per dollar of compensation. In public K–12 education, by contrast, the compensation “system” is fragmented and uncoordinated, with provisions often determined by means not based on systematic assessment of the overall incentive effects, such as pressures from a particular constituency or inherited from earlier contracts.

Accountability pressures are forcing school districts to address the inefficiencies built into this compensation system, and rethink how they are spending roughly \$250 billion annually for compensation of instructional personnel. Federal programs such as the Teacher Incentive Fund are encouraging states to experiment with performance-related and market-driven pay. States such as Minnesota, Florida, and Texas have developed programs to encourage their school districts to move towards performance-related and market-driven pay structures. Texas has taken steps in this direction with the Texas Educator Excellence Grant (Springer et al., 2009) and, more recently, the DATE program (Springer et al., 2010a).

A number of large urban districts have also taken important steps in this direction. Performance-related and market-driven pay incentives are much more common in charter schools and are expanding with the charter school base (Podgursky, 2009). Taken as a whole, performance-related pay slowly seems to be working its way into the teacher compensation landscape. That said, school administrators should move forward thoughtfully, and make use of best use of education research, experience from other sectors, as well as principles of personnel economics, particularly as the slender base of U.S. education research expands.

Policy makers and education stakeholders at all levels would benefit from rigorous assessments of teacher compensation reform programs and policies, as well as assessments of the effect of their various design components. For instance, should individual teachers or teams of teachers be rewarded, or perhaps a combination of both? Should the measure be based on student growth or attainment? What criteria should be included? Should it be based strictly on student test scores, or should other measures, like principal evaluations, be included? If other measures are included, what should be the weight of each element? We are in no position at present to specify what, if any, "best practices" exists for performance-related pay systems. Indeed, it may be the case that there exists no single best plan, but rather a menu that can be customized to fit the peculiarities of local labor markets and school organizations.

Given the relatively negative findings for the United States, one might conclude performance-related pay programs tied to classroom or school-wide achievement gains are not a useful avenue for school reform. However, it should also be noted that these evaluation studies are relatively short term and focused on the effect of incentive programs on the current teaching workforce. In the long term incentive pay programs may shape the workforce in positive ways, if more effective (and thus higher paid) teachers are more likely to remain on the job or enter the applicant pool, and less effective teachers leave. We noted in our survey of the literature that one repeated finding is that some teachers are consistently more effective than others in producing student achievement gains and that the difference between the top and bottom deciles of teachers is educationally substantial. What the U.S. findings suggest is that, thus far, these incentive plans have not been very successful in "growing" more effective teachers. However, given the wide range of teacher effectiveness in districts and schools, it may be that the biggest gains can come from personnel policies (including compensation) that aim to change composition of the teacher workforce.

By depending so heavily on the single salary schedule, U.S. public school systems have not been able to leverage incentives to retain high performers and shed low performers. A more efficient pay structure would focus on retaining the best teachers while pushing out those instructors not meeting expectations. Indeed, it is well recognized in the personnel economics and general management literatures that differential recruitment and retention of more productive employees can be at least as important as performance gains attributed to the motivational response among workers (Podgursky and Springer, 2007; Lazear and Polachek, 2005; Lazear, 2000, 2003).

As schools districts continue to experiment with alternatives or supplements to traditional salary schedules, hopefully more districts and schools will design and implement interventions in ways that permit rigorous and long-term evaluation. As that happens, we will get better insight as to effective alternatives to traditional salary schedules. While researchers need to also begin exploring the cost-effectiveness of incentive pay systems, it is equally important to consider redesign of other components of educator compensation systems, including salary structures, pay grades within discrete job bands, and retirement benefits.

ACKNOWLEDGMENTS

This paper was prepared for the Harvard Program on Education Policy and Governance conference on merit pay, Cambridge, Massachusetts, June 2–4, 2010. The authors wish to acknowledge two federally-funded national research and development centers for research support — the National Center on Performance Incentives at Vanderbilt University and the Center for Analysis of Longitudinal Data in Education Research. We also thank Yi Du for research assistance. This paper also benefitted from helpful comments and feedback by Dale Ballou, John Diamond, Cate Gardner, Steve Glazerman, and Michael Henderson. The usual disclaimers apply.

REFERENCES

- Aaronson, Daniel, Lisa Barrow, and William Sanders, 2007. "Teachers and Student Achievement in Chicago High Schools." *Journal of Labor Economics* 25 (1), 95–136.
- Bacolod, Marigee, 2007. "Do Alternative Opportunities Matter? The Role of Female Labor Markets in the Decline of Teacher Quality." *Review of Economics and Statistics* 89 (4), 737–751.
- Ballou, Dale, 2001. "Pay for Performance in Public and Private Schools." *Economics of Education Review* 20 (1), 51–61.
- Ballou, Dale, and Michael J. Podgursky, 1997. *Teacher Pay and Teacher Quality*. W.E. Upjohn Institute for Employment Research, Kalamazoo, MI.
- Barrow, Josh, and Stuart Buck, 2010. *Underfunded Teacher Pension Plans: It's Worse Than You Think*. Civic Report No. 61. Manhattan Institute for Policy Research, New York, NY.
- Clark, Robert L., 2009. "Retiree Health Plans for Public School Teachers After GASB 43 and 45." *Education Finance and Policy* 5 (4), 438–462.
- Clotfelter, Charles, Elizabeth Glennie, Helen Ladd, and Jacob Vigdor, 2008. "Teacher Bonuses and Teacher Retention in Low Performing Schools: Evidence from the North Carolina \$1,800 Teacher Bonus Program." *Public Finance Quarterly* 36 (1), 63–87.
- Clotfelter, Charles, Helen Ladd, and Jacob Vigdor, 2006. "Teacher-Student Matching and the Assessment of Teacher Effectiveness." *Journal of Human Resources* 41 (4), 778–820.

- Corcoran, Sean P., William N. Evans, and Robert M. Schwab, 2004. "Changing Labor-Market Opportunities for Women and the Quality of Teachers, 1957-2000." *American Economic Review* 94 (2), 230-235.
- Costrell, Robert M., and Michael J. Podgursky, 2009. "Peaks, Cliffs, and Valleys: The Peculiar Incentives of Teacher Retirement Systems and their Consequences for School Staffing." *Education Finance and Policy* 4 (2), 175-211.
- Ehrenberg, Ronald G., 2004. "Prospects in the Academic Labor Market for Economists." *Journal of Economic Perspectives* 18 (2), 227-238.
- Figlio, David A. and Lawrence Kenny, 2007. "Individual Teacher Incentives and Student Performance." *Journal of Public Economics* 91 (5-6), 901-914.
- Folland, Sherman, Allen C. Goodman, and Miron Stano, 2006. *The Economics of Health and Health Care*, 5th Edition. Prentice-Hall, Upper Saddle River, NJ.
- Glazerman, Steven, Allison McKie, and Nancy Carey, 2009. *Evaluation of the Teacher Advancement Program (TAP) in the Chicago Public Schools: Study Design Report*. Document No. PR08-14. Mathematica Policy Research, Inc., Washington, DC.
- Glewwe, Paul, Nauman Ilias, and Michael Kremer, 2008. *Teacher Incentives*. Unpublished manuscript. Harvard University, Cambridge, MA.
- Goldhaber, Dan, Hyung-Jai Chung, Michael DeArmond, and Daniel Player, 2005. *Why Do So Few Public School Districts Use Merit Pay?* Paper presented at APPAM Fall Research Conference, November 3-5. Washington, DC.
- Goldhaber, Dan, and Michael Hansen, 2010. "Using Performance on the Job to Inform Teacher Tenure Decisions." *American Economic Review* 100 (2), 250-255.
- Goodman, Serena, and Lesley Turner, 2011. *Teacher Incentive Pay and Educational Outcomes: Evidence from the NYC Bonus Program*. Paper presented at Harvard Kennedy School Conference on Merit Pay, June 2-4, Cambridge, MA.
- Grogger, Jeff, and Eric Eide, 1995. "Changes in College Skills and the Rise in the College Wage Premium." *Journal of Human Resources* 30 (2), 280-310.
- Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw, 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica*, 69 (1), 201-209.
- Hanushek, Eric, 2003. "The Failure of Input-based Resource Policies." *Economic Journal* 113 (485), F64-F68.
- Hanushek, Eric, 2009. "Teacher Deselection." In D. Goldhaber and J. Hannaway. (eds.), *Creating a New Teaching Profession*, 165-180. Urban Institute Press, Washington, DC.
- Hanushek, Eric, John Kain, Daniel O'Brien, and Steven Rivkin, 2005. "The Market for Teacher Quality." NBER Working Paper 11154. National Bureau of Economic Research, Cambridge, MA.

- Hanushek, Eric, and Steven Rivkin, 2004. "How to Improve the Supply of High-Quality Teachers." *Brookings Papers on Education Policy*, 7-25. Brookings Institution, Washington, DC.
- Hudson, Sally, 2010. *The Effect of Performance-Based Teacher Pay on Student Achievement*. Discussion Paper 09-023. Stanford Institute for Economic Policy Research, Palo Alto, CA.
- Iaterola, Patrice, and Leanna Steifel, 2003. "Intradistrict Equity and Public Education Resources and Reform." *Economics of Education Review* 22 (1), 60-78.
- Ingersoll, Richard M., 1999. "The Problem of Underqualified Teachers in American Secondary Schools." *Educational Researcher* 28 (2), 26-37.
- Kane, Thomas, Jonah Rockoff, and Douglas Staiger, 2008. "What Does Certification Tell Us About Teacher Effectiveness? Evidence from New York City." *Economics of Education Review* 27 (6), 615-631.
- Kershaw, Joseph, and Roland McKean, 1962. *Teacher Shortages and Salary Schedules*. RAND Document RM-3009-FF. RAND Corporation, Santa Monica, CA.
- Kirshtein, Rita J., Andrea R. Berger, Elana Benatar, and David Rhodes, 2004. *Workforce Contingent Financial Aid: How States Link Financial Aid to Employment*. Lumina Foundation, Indianapolis, IN.
- Lankford, Hamilton, Susanna Loeb, and James Wyckoff, 2002. "Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis." *Educational Evaluation and Policy Analysis* 24 (1), 37-62.
- Lavy, Victor, 2002. "Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement." *Journal of Political Economy* 110 (6), 1286-1317.
- Lavy, Victor, 2009. "Performance Pay and Teachers' Effort, Productivity, and Grading Ethics." *American Economic Review* 99 (5), 1979-2011.
- Lazear, Edward P., 2000. "Performance Pay and Productivity." *American Economic Review* 90 (5), 1346-1361.
- Lazear, Edward P., 2003. "Teacher Incentives." *Swedish Economic Policy Review* 10 (2), 179-214.
- Lazear, Edward P., and Solomon W. Polacheck, 2005. "Output-based Pay: Incentives or Sorting." In Polacheck, Solomon. (ed.), *Research in Labor Economics, Volume 23, Accounting for Worker Well-Being*, 1-25. Emerald, New Milford, CT.
- Lee, David, and Thomas Lemieux, 2009. "Regression Discontinuity Designs in Economics." NBER Working Paper No. 14723. National Bureau of Economic Research, Cambridge, MA.
- Levin, Jessica, Jennifer Mulhern, and Joan Schunck, 2005. *Unintended Consequences: The Case for Reforming the Staffing Rules in Urban Teachers Union Contracts*. The New Teacher Project, New York, NY.

Springer, Matthew G., and Marcus Winters, 2009. *New York City's School-Wide Bonus Pay Program: Early Evidence from a Randomized Trial*. NCPI Working Paper 2009-02, National Center on Performance Incentives, Nashville, TN.

Steele, Jennifer, Richard J. Murnane, and John B. Willett, 2010. "Do Financial Incentives Help Low-Performing Schools Attract and Keep Academically Talented Teachers? Evidence from California." *Journal of Policy Analysis and Management* 29 (3), 451–478.

Taylor, Lori L., and Matthew G. Springer, 2010. *Designing Incentives for Public Sector Workers. Evidence from the Governor's Educator Excellence Grant Program in Texas*. NCPI Working Paper. National Center on Performance Incentives, Nashville, TN.

Thistlewaite, Donald, and Donald Campbell, 1960. "Regression Discontinuity Analysis: An Alternative to the Ex-Post Factor Experiment." *Journal of Education Psychology* 51 (6), 309–317.

United States Department of Education, 2009. *Digest of Education Statistics*. U.S. Department of Education, National Center for Education Statistics, Washington, DC.

Varadarajan, Poondi, and Charles Futrell, 1984. "Factors Affecting Perceptions of Smallest Meaningful Pay Increases." *Industrial Relations* 23 (2), 278–286.

Copyright of National Tax Journal is the property of National Tax Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

