Show *Who* the Money? Does Performance Pay Attract Higher Quality Teachers?[1]

Michael Jones
Research Assistant Professor
Department of Economics
323 Lindner Hall
University of Cincinnati
Cincinnati, OH 45221
513-556-2491 (office)
m.jones@uc.edu

Michael T. Hartney
Ph.D. Candidate
Department of Political Science
217 O'Shaughnessy Hall
University of Notre Dame
Notre Dame, IN 46445
561-644-2523 (office)
mhartney@nd.edu

Abstract: Advocates of teacher pay-for-performance often claim that redesigning teacher compensation systems to include performance incentives will reshape entry into the profession leading to a higher quality teaching force in the long run. However, the link between performance pay and teacher selection is largely untested in the research literature on teacher compensation. Drawing on two waves of the Schools and Staffing Survey (SASS) we find that, on average, districts that implemented performance pay to reward excellence in teaching secured new teacher hires who graduated from colleges and universities with average incoming SAT scores that were about 30 points higher than the new teacher cohorts hired by districts that did not adopt performance pay programs. These findings are robust to a variety of model specifications and persist even when we control for districts' use other pay recruitment incentives including: market-pay, hard-to-staff schools pay, and more generous baseline salaries.

Keywords: Compensation, Pay Reform, Performance Pay, Teacher Recruitment, Teacher Quality

[1] Previous versions of this article were presented at the 2012 meeting of the Association for Education Policy and Finance in Boston, MA and at the 2012 meeting of the Association for Public Policy Analysis & Management in Baltimore, MD. We thank Jacob Vigdor, Kristine Lamm West, Alex Bowers, Daniel Bowen, Jonathan Mills, Nathan Jensen, Gary Ritter, and Matthew Springer for their feedback. We also thank David E. Campbell, Jill Constantine, William N. Evans, Michael Hansen, and participants at the University of Notre Dame's Rooney Center for the Study of American Democracy workshop for helpful comments and suggestions. All errors are, of course, our own.

1. Introduction

> "Money is never the reason why people enter teaching, but it is the reason why *some* people do not enter teaching…"
>
> –Arne Duncan, U.S. Education Secretary (2011)

> "The trustees [of New York City's Board of Education in the 1830s] expected their low-paid teachers to be paragons of virtue and promised that while they might or might not receive their 'merited reward' from men, they would certainly receive 'what is infinitely more valuable—the approbation of Heaven.'"
>
> –Diane Ravitch, *The Great School Wars* (1974)

Today, education policy-makers have forsaken relying on divine providence to deliver unto them an effective teacher workforce for the 21st century. Instead, the public officials who govern America's schools are in search of evidence-based policies that will help public education attract top talent into the teaching profession (Berry 1988; Goldhaber and Hannaway 2009).[2] Policy-makers' laser-like focus on recruiting better teachers is based on more than good intuition. In the nearly forty years since James Coleman (1966) released his groundbreaking report, *Equality of Educational Opportunity*, the most reliably consistent finding in education research is that the quality of a child's classroom teacher is the single most important school-level variable influencing student achievement (Aaronson, Barrow and, Sander 2007; Chetty, Friedman, and, Rockoff 2011; Rivkin, Hanushek, and, Kain 2005; Sanders and Rivers 1996).[3]

---

[2] The significant decline in the percentage of high aptitude candidates (particularly women) entering the teaching profession since mid-century is well-chronicled by Corcoran, Evans, and, Schwab (2004a, 2004b).

[3] For example, Hanushek and Rivkin (2004, p. 21) summarize this literature by noting that, "having five years of good teachers in a row (one standard deviation above average) could overcome the average seventh grade mathematics achievement gap between lower income kids and those from higher income families." More recently, Chetty, Friedman, and, Rockoff (2011) conclude that, "Students assigned to [highly effective]

This newfound consensus that improving teacher quality is the key to increasing student achievement has led some policy-makers to ask whether changing the way that teachers are compensated could help raise the overall quality of the nation's teaching force. Among the most hotly debated pay reform proposals are those that propose scrapping public education's traditional salary schedule that compensates teachers primarily for years of experience and degrees accrued for a performance-based system that instead ties educator pay to actual measures of teaching effectiveness. Such proposals have drawn widespread interest across the American political system—from big city mayors active in Chicago, Denver, and New York to national political figures, including presidents in both political parties.[4] And, at the state level, governors— long concerned with the linkage between economic competitiveness and the quality of public education—have helped enact statewide performance pay policies that offer financial bonuses to teachers whose students achieve significant learning gains (e.g., Minnesota, Florida, and Texas). Figure 1 (below) provides a visual perspective of the varied use of performance pay across the U.S. in the 2007-08 school year. As can be seen in Figure 1, the percentage of public school districts using performance pay varies from a high of 49 percent of Florida's school districts to fewer than 1 percent of districts in Washington and Nebraska.

---

teachers are more likely to attend college, attend higher-ranked colleges, earn higher salaries, live in higher SES neighborhoods, and save more for retirement."

[4] In support of this claim, consider the fact that the last two presidential administrations have enacted major education reforms – President Obama's Race to the Top program and President Bush's Teacher Incentive Fund program – that devote significant attention to reforming the way K-12 teachers are recruited, compensated, and evaluated.

[Figure 1 about here]

Predictably, these efforts to expand the use of performance pay in public education have been met with great controversy (Buck and Greene 2011; Goldhaber et al. 2008). Opponents seize on a number of studies that raise questions about the reliability of the value-added models that reformers propose to use to assess teacher effectiveness (McCaffery et al. 2004; Rothstein 2009; Rothstein 2010; but see Koedel and Betts 2011). In response, advocates of teacher pay-for-performance ask why teachers ought to be uniquely shielded from the type of pay policies that routinely govern employer-employee relations in other work environments where managers' holistic evaluations are equally imperfect at gauging employee effectiveness (Hess 2004). Notably, advocates of teacher pay-for-performance advance another important set of claims about the potential benefits of linking teacher compensation to teacher effectiveness. In short, they argue that the introduction of performance pay will have powerful *selection effects* on the teaching profession itself, potentially attracting a different caliber candidate into teaching—individuals who might not otherwise consider a career in the public education workforce (Anderson quoting Hanushek 2010; Leigh 2013).[5]

Because this entire claim—a major argument advanced by those who support greater experimentation with teacher pay-for-performance— is largely untested in the research literature on teacher compensation, we focus on it here by addressing the potential of performance pay to influence teacher selection. Specifically, we ask: does

[5] As Hanushek (2010) argues, "the biggest role of incentives [paying teachers based on performance] has to do with selection of who enters and who stays in teaching – i.e., how incentives change the teaching corps through entrance and exits…" (Anderson quoting Hanushek, 2010, p. A18).

teacher performance pay appeal to a certain *type* of teacher? To answer this question, we use data from a nationally representative survey of teachers and public school districts and conduct a difference-in-difference analysis comparing the quality of teachers selecting into school districts that adopted teacher pay-for-performance between 2003-04 and 2007-08. We find that, on average, public school districts that adopted performance pay programs were significantly more successful in luring higher quality candidates into their school districts than an otherwise similar set of districts that retained traditional teacher pay systems. Specifically, school districts that implemented performance pay to reward excellence in teaching secured new teachers who graduated from colleges and universities with incoming freshman SAT scores that were about 30 points higher than those teachers who took positions in otherwise similar school districts that did not offer performance pay.

The remainder of this article is organized as follows. In the next section, we review the existing literature on teacher performance pay and situate our own study among this prior work. In the third section of the paper, we describe the data used in our analyses. Section four explains our identification strategy. We present the results of our analyses in section five. Finally, we conclude with a discussion of the results including key limitations of the study, implications for policy-makers, and promising avenues for future research.

2. Existing Literature

Identifying the effects of performance pay on teacher selection allows us to address a broader question in labor economics about the role of selection effects in "tournament-style" compensation, while simultaneously answering a policy relevant question about

whether performance pay could help improve the overall quality of teachers recruited into U.S. school districts.

Outside the domain of public education, labor economists have already identified powerful sorting effects in "tournament" style compensation systems. Depending on the circumstance and task, these studies suggest that the selection effects of performance pay likely trump the basic incentive effects of such programs (Eriksson et al. 2009; Lazear 2000; Leuven et al. 2011). For example, Leuven et al. (2011) conduct a randomized field experiment to disentangle selection/sorting effects from incentive-based ones in tournament pay environments and find causal evidence that "[tournament participants] of higher ability are more likely to select themselves into tournaments with higher rewards" (Leuven et al. 2012, p. 654). To estimate the portion of performance pay effects that accrue solely on the basis of higher-ability individuals exhibiting a preference for working in performance-oriented settings, Eriksson et al. (2009) performed an experiment in which they allowed subjects to self-select into either fixed or variable-based compensation systems. Consistent with Leuven et al.'s study, their experiment likewise revealed that performance pay exhibited both an "effort" and a "sorting/selection" effect on participants with the sorting effect "[making] up a little more than half of the total increase in effort…." (Eriksson et al. 2009, p. 544).

While the research literature on teacher compensation contains numerous studies examining the effects of teacher performance pay (Atkinson et al. 2009; Figlio and Kenny 2007; Goldhaber and Walch 2012; Imberman and Lovenheim 2012; Lavy 2002, 2009; Winters et al. 2008; Woessmann 2011), including many randomized experiments (Fryer et al. 2013; Fryer, Forthcoming; Glewwe, Ilias, and, Kremer 2010; Muralidharan and

Sundararaman 2011; Springer et al. 2010), nearly all of them focus exclusively on whether financial incentives motivate current veteran teachers to exert greater effort and in turn raise their students' test scores.[6] For example, Springer et al.'s (2010) multi-year random-assignment POINT study estimated the causal effects of offering teacher performance pay on teachers' value-added effectiveness by examining whether a randomly selected "treatment" group of teachers eligible for a $20,000 performance bonus demonstrated higher subsequent value-added scores than an otherwise identical "control" group of teachers ineligible for the award. While POINT garnered widespread attention for finding no relationship between performance bonuses and higher value-added teacher performance, the study was never designed to assess whether continued use of performance pay in school districts like MNPS could, over the long run, improve the composition of its teaching force by attracting more accomplished candidates into the district (Sawchuck 2010). This shortcoming in the existing literature on teacher compensation is well articulated by Lazear (2003, p. 186) who notes that, "Most of the discussion on performance or output-based pay centers on incentives. Relatively little attention is given to the nature of people that various compensation schemes attract, but selection is equally important and in the context of education, perhaps the most important aspect of compensation effects."

To remedy scholarly inattention to the selection side of teacher pay reform, a handful of recent studies have examined the relationship between teacher compensation and sorting preferences within the profession (Bowen and Mills 2012; Bowen et al. 2013; Hendricks 2011; Perez, Muraki, and Loeb 2011). Perez, Muraki, and Loeb (2011) make a

[6] For a comprehensive review of this research literature see Podgursky and Springer (2007).

particularly notable contribution by linking administrative data from Florida to an original survey of teachers that asked each educator about the design of her preferred compensation system. Perez and her colleagues found that "highly effective teachers (as measured by value-added) were more likely to select an individual competitive bonus over a salary increase" as their preferred system of teacher compensation (Perez, Muraki, and Loeb 2011, p. 26). In another study, Hendricks (2011) examined whether teachers who self-selected into performance pay districts were more likely to receive an actual performance bonus; however, he found no evidence that those teachers who self-selected into these districts did in fact earn a performance pay bonus. While an important contribution, Hendricks' findings could also be attributed to flaws in the implementation of the performance pay program analyzed, rather than reflecting the quality of incoming teachers across a representative sample of performance pay districts. Finally, in two separate studies Bowen and colleagues (2012, 2013) use a series of laboratory experiments to demonstrate that teachers are generally more risk-averse than comparable non-teacher professionals and that the most risk-averse teachers prefer working under fixed (e.g. rigid salary schedules), versus variable pay (e.g. pay-for-performance) systems.

Related survey-based research suggests that more academically accomplished college graduates prefer compensation systems that are structured, in part, on the basis of an employee's performance. Specifically, surveys of undergraduate seniors attending America's most highly ranked colleges and universities recently revealed that top graduates would not seriously consider pursuing careers in teaching unless a dramatic change in the compensation structure of public education were to occur— including more frontloaded salaries and less rigid wage frictions that allow for pay to be differentiated by

performance (Auguste, Kihn, and, Miller 2010). Specifically, on the basis of the answers provided by these survey respondents it was estimated that, "offering a 20% performance bonus to the top performing 10% of teachers would induce roughly an 11% increase in the number of top-third students becoming teachers" (Auguste, Kihn, and, Miller 2010, p. 31). In sum, while these studies all offer useful insights into the *attitudes* prospective and current teachers hold toward performance pay, our study represents an effort to move beyond the laboratory setting and examine whether performance pay policies influence teacher selection *behavior* in and across a representative sample of U.S. school districts.

3. Data

In order to test the hypothesis that performance pay affects teacher selection by appealing to higher quality prospective teaching candidates, we need two types of data. First, we need data identifying variation in teacher compensation policies across a representative sample of U.S. school districts. Second, we need a defensible way to measure the quality of the incoming teacher cohorts hired by those same districts both before and after the district's introduction of performance pay. One data source that does a good, though admittedly imperfect job, of providing these two complementary measures is the Schools and Staffing Survey (SASS) restricted-use data files.[7] Using the two most recent cycles of SASS (2003-04, 2007-08) we construct a short district-level panel data set that contains information about each district's use of differential pay policies (including

[7] Begun in 1987 and administered by the National Center for Education Statistics (NCES), the SASS is fielded every three to four years and surveys a randomly selected sample of public schools, private schools, and schools funded by the Bureau of Indian Education (BIE). The SASS collects data from district- and school-level administrators, principals, and classroom teachers.

performance pay) as well as information about the individual-level characteristics of the teachers working in those same districts.

3.1 Measurement

Both the 2003 and 2007 waves of the SASS asked each sampled school district whether it used "any pay incentives such as cash bonuses, salary increases, or different steps on the salary schedule to reward excellence in teaching." We use a district's response to this dichotomous survey item (1=yes; 0=no) to classify whether that district offered some type of teacher performance pay in the given school year. Unfortunately, SASS offers no other details on the size and scope of the bonus awards offered to teachers under its performance pay plan.[8]

In order to focus on the effects of our key explanatory variable of interest—whether a district offers performance pay—we need to account for other district-level characteristics that are likely to influence where teachers elect to work. For example, a district's overall academic performance is one obvious factor that could impact teacher selection since the quality of the students that teachers will be assigned determines the relative difficulty/desirability of the position assignment. To account for cross-district variation in academic performance, we gathered estimates of the high school graduation rate for each district in the SASS sample drawing on *Education Week's* District Graduation

[8] This measure is especially limiting in that it does not allow us to examine potentially important differences in how teacher selection is shaped by variation in the type of performance pay program a district uses (e.g. group-based versus individual incentives, large versus small awards, high versus low qualification thresholds). We elaborate on this shortcoming when we discuss the study's overall limitations and the policy implications of our findings in the paper's concluding section.

Rate Map Tool.[9] Using the Common Core of Data (CCD), *Education Week* estimates a district's graduation rate using the cumulative promotion index. The following formula is used to calculate each school district's 2007 graduation rate -

$$\frac{10^{th}\text{ graders, fall 2007}}{9^{th}\text{ graders, fall 2006}} \times \frac{11^{th}\text{ graders, fall 2007}}{10^{th}\text{ graders, fall 2006}} \times \frac{10^{th}\text{ graders, fall 2007}}{11^{th}\text{ graders, fall 2006}} \times \frac{\text{HS graduates, spring 2007}}{12^{th}\text{ graders, fall 2006}}$$

To further account for major differences in the types of school districts that reported using performance pay in the 2007-08 school year, we examined both district expenditures and graduation rates four years prior to the implementation of performance pay. Using the same *Education Week* graduation rate estimates and district-level expenditures derived from the CCD, we assigned one of four possible classifications to each school district in the SASS sample based on whether the district posted *above/below* average graduation rates and reported *above/below* average per pupil expenditures.[10] We then use where a district fits in this classification scheme as an added control variable in our subsequent empirical analyses to help capture a district's relationship between expenditures and graduation rates, beyond controlling for each one separately. Following other work on teacher labor markets, we also include controls for district size (student enrollment), district locale

[9] The SASS public school district survey also asks each district official for an estimate of the district's graduation rate, but given the latitude with which states have historically reported their graduation rates we elected to use the *Education Week* measure because it uses the same approach to estimate rates across states.

[10] The four possible combinations are: High Maintenance District (above average expenditures and above average graduation rates), Low Maintenance District (below average expenditures and below average graduation rates), Overachieving District (below average expenditures and above average graduation rates), and Underachieving District (above average expenditures and below average graduation rates).

(urban vs. rural), district working conditions (teacher-student ratio), student and teacher demographics (student poverty, percentage of non-white students and teachers), and whether or not a district offered other (non-performance pay) teacher pay incentives that might influence teacher selection decisions (a generous baseline salary schedule, "market-based" pay, "hard-to-staff schools" pay, and additional pay for earning National Board Certification).

Consistent with prior studies that detect a limited use of performance pay in public education (Ballou and Podgursky 1997; Ballou 2001; Goldhaber et al. 2008), we similarly find that a mere nine percent of the districts SASS sampled reported using performance pay in the 2007-08 school year. Table 1 (below) provides a descriptive look at how school districts that indicated offering performance pay differed from districts that did not offer such incentives.

[Table 1 about here]

Notably, performance pay districts are significantly more likely to be located in urban areas, employ a lower percentage of white teachers, and serve a higher percentage of students who qualify for free or reduced lunch. Only forty percent of performance pay districts reported having a collective bargaining agreement in place compared to sixty percent of non-performance pay districts. Performance pay districts also tend to graduate fewer of their students on time from high school and have higher student-teacher ratios. Significantly more performance pay districts are classified as "low maintenance" districts under our district-type classification scheme. That is, performance pay districts were far more likely to report expenditures and graduation rates that were well below the average in the SASS sample. The use of performance pay in these districts may indicate an attempt

by district policy-makers to use performance pay bonuses as a low-cost way of increasing student achievement. We also find that performance pay districts are much more likely to offer other (non-performance based) types of financial incentives to teachers who work in less desirable schools as well as to teachers who teach in a field experiencing a labor market shortage.

To construct our dependent variable—our proxy measure of a prospective candidate's quality prior to their actual entry into the profession—we use the information SASS collects in its teacher survey identifying the postsecondary institution where each teacher received her undergraduate degree. We then use the average aptitude-based selectivity requirements (SAT scores) of each teacher's undergraduate institution as a proxy to measure the quality of an individual teacher's bachelor's degree and by extension a related proxy capturing a teacher's academic aptitude prior to their entering the teacher labor market. We fully acknowledge that our approach to measuring the anticipated quality of *prospective* teaching candidates is less than ideal; however, our approach to measurement follows a long tradition in the teacher quality literature of using measures related to an individual's academic aptitude as a proxy for teacher quality when actual measures of classroom effectiveness are not available (Ehrenberg and Brewer 1994; Figlio 1997; Player 2009). Moreover, even though a teacher's past effectiveness is the most reliable predictor of their future performance (Goldhaber 2002), academic aptitude has, for the most part, remained one of the few observable characteristics that scholars have been able to link to teacher effectiveness. For example, Rockoff, Jacob, Kane, and Staiger (2008) found a positive relationship between college selectivity and teacher effectiveness for New York City teachers. Similarly, Clotfelter, Ladd, and, Vigdor (2007) found that "a teacher who

comes from an undergraduate institution ranked as competitive appears to be somewhat more effective on average than one from an uncompetitive institution." And, Ehrenberg and Brewer (1994) show that the selective quality of a teacher's postsecondary degree institution was significantly related to the learning outcomes of their students in the neighborhood of a 2-percentage point increase on student test scores. In sum, Hanushek's (1981) two decades old summation of the research literature on teacher quality is for the most part as equally true today as it was when he first wrote that, "The only reasonably consistent finding seems to be that 'smarter' teachers do better in terms of student achievement" (Hanushek p. 29, 1981).

3.2 Sample Restrictions

We restrict all of our analyses to the 2003 and 2007 SASS waves. Because we are interested in examining whether a school district's adoption of performance pay *differentially* influences high-aptitude teachers to seek and obtain employment in these school districts, we required data with sufficient observations at two discrete points in time. Our decision to examine the selection behavior of teachers with fewer than four years of experience in the 2007-08 wave of SASS is an especially important component of our research design. This restriction further allows us to focus on the potential effect offering performance pay may have on the selection behavior of high-aptitude early-career teachers. That is, we emphasize whether performance pay distinctly appeals to the quality of a school district's new (incoming) teacher hires in 2007 by restricting our analyses to the hiring that took place between 2003 and 2007 looking only at the sorting behavior of teachers with fewer than four years experience.

4. Identification Strategy

4.1 OLS

Beginning with an OLS regression model, we write our baseline estimating equation as the following -

$$Y_{ds} = \beta_0 + PerformancePayDistrict_{ds}\beta_1 + X_{ds}\beta_2 + u_s + \varepsilon_{ds} \quad (1)$$

where Y is a measure of new teacher selectivity for school district d in state s, reflecting newly hired teacher respondents in the 2007-08 SASS. Note that the level of observation in this analysis is the school district, not the individual teacher, and the data is weighted by the inverse of the probability of selection (sample weights are provided by SASS). We use several different teacher selectivity indicators (dependent variables) to proxy the overall quality of each district's newly hired teacher cohort including: the average undergraduate institution SAT median, SAT 75th percentile ranking, and SAT 25th percentile ranking. β_1 is our coefficient of interest in equation 1 since it measures how the average quality of new teachers changes in districts where performance pay has been implemented. X_{ds} is a vector of district covariates including: the percentage of students eligible for a free lunch, student-teacher ratio, the presence of a collective bargaining agreement, a district's use of other teacher policy reforms (e.g. recruitment incentives), and basic demographic characteristics of the district's teachers. We further include state fixed effects and ε_{ds} represents the idiosyncratic error term.

OLS estimates of equation 1 are, however, unlikely to identify the true causal effect of a district's adoption of performance pay on the quality of new teachers selecting into the district if school districts that enact performance pay policies have underlying

characteristics that are fundamentally different from districts which do not adopt performance pay. Both theory and prior research jointly suggest that school districts that adopt performance pay policies—policies which are regarded as controversial—differ significantly from school districts that rely rigidly and exclusively on a traditional single-salary schedule to compensate their teachers. For example, the enactment and persistence of district-level performance pay programs has been correlated with cross-district variation in: teacher union strength (Ballou 2001; West and Mykerezi 2011), political climate (Goldhaber et al. 2008), and fiscal health/budgetary resources (Leigh 2013). For all of these reasons, we must be careful to heed the advice of Figlio (2002, p. 687) who has cautioned scholars analyzing the relationship between teacher compensation and teacher selection that, "[disentangling] the effects of the self-selection of teachers for non-salary reasons into high-paying districts from the effects of salary as a means of recruiting better-qualified teachers…requires more information than a cross section salary regression can provide." Therefore, even with the rich set of control variables included in our OLS estimates, we anticipate the existence of bias on account of unobserved differences that exist between performance pay versus non-performance pay districts. Due to these endogeneity concerns under the basic cross-section OLS research design, we further test our hypothesis that performance pay shapes teacher selection by using an additional, more rigorous empirical test: a difference-in-difference estimator that can account for unobserved time invariant district-level characteristics.

4.2 Difference-in-Difference

Using the two most recent cycles of SASS (2003-04, 2007-08) we construct a short district-level panel data set that contains information about each district's use of

differential pay policies (including performance pay) as well as information about the individual-level characteristics of the teachers working in those same districts. We restrict the panel sample to those districts that did not use any pay incentives to recruit teachers in 2003-2004. We further drop from our difference-in-difference analysis any school districts that did not appear in both the 2003 and 2007 waves of the SASS as we cannot observe those districts' compensation policies in both years. Our "treatment" group in the difference-in-difference analysis consists of those school districts that implemented performance pay sometime between 2003 and 2007. The control group are districts that had not implemented performance pay by 2007-08. The identifying assumption behind our difference-in-difference estimation strategy is that the trend in the quality of teacher recruitment for the treatment group would be the same as the control group in the absence of performance pay. We test this assumption as a robustness check later in the paper and find that the assumption does indeed hold. We can then estimate the following equation –

$$Y_{dst} = \beta_0 + 2007_{dst}\beta_1 + PerformancePay_{dst}\beta_2 + 2007*PerformancePay_{dst}\beta_3 + X_{dst}\beta_4 + \mu_s + \varepsilon_{dst} \quad (2)$$

where Y_{dst} is a measure of new teacher selectivity for school district d in state s at time t, 2007 is a dummy variable for the year 2007, $PerformancePay_{dst}$ is a dummy variable indicating if a district implemented performance pay, $2007*PerformancePay_{dst}$ is an interaction term and the variable of interest in the difference-in-difference model, X_{dst} is the same vector of district characteristics in the OLS estimating equation 1, μ_s are state effects, and ε_{dst} is an idiosyncratic error term.

We calculate standard errors by using replicate weights because the SASS is a complex survey design (i.e. samples are stratified, clustered, and weighted). Replicate

weights are used to generate more accurate standard error estimates by retaining all the information about the complex sample design. Standard errors are calculated after constructing subsamples, or replicates, from the full sample. Instead of creating replicates by randomly drawing observations (as in some bootstrapping methods), the method of replicate weights incorporates the stratification and clustering information of the SASS. The difference between the point estimate in the full sample and the point estimate using the weights of each of the replicates is used to determine the standard error.[11]

5. Results

5.1 OLS

We first present results from the OLS regression analysis. Table 2 shows the results of regressing the quality of new teachers hired by the sample of SASS school districts in 2007 on a district's performance pay policy (1=yes performance pay used; 0=no, performance pay not used) in addition to the basic demographic and district-level control variables reported in the methodology section above.

[Table 2 about here]

Our OLS estimates show that, compared to the quality of new teacher cohorts in non-performance pay districts, new teacher cohorts in performance pay districts earned their baccalaureate degrees from institutions with SAT scores that are 17 points higher on average. This result is robust to using either the 75th percentile or the 25th percentile of the postsecondary institution's average SAT score. This 17 point difference corresponds

[11] More details on the NCES recommendation to use replicate weights for standard errors, as well as their construction, can be found at http://nces.ed.gov/pubs98/98312.pdf

substantively to an increase of 0.075 standard deviations on SAT scores, and a ranking approximately 40 places higher on the complete list of 1431 colleges and universities ranked by incoming freshmen students' SAT scores.[12]

Although some hesitate to place much importance on the average SAT scores for a postsecondary institution's income freshman class, the average SAT of accepted students at a college or university is highly correlated with the prestige and status of postsecondary institutions, particularly when it comes to the perception of the mainstream public. We argue that one potential consequence of this "prestige perception" is a conferral of additional status on graduates from these institutions—where graduates come with a premium in the labor market after they graduate from a prestigious college (see also, Brewer, Eide, and, Ehrenberg 1999). The fact that we find graduates from more prestigious colleges and universities selecting into performance pay districts at rates higher than average suggests that performance pay may be one policy lever available to school districts that wish to attract a larger share of their future teacher cohorts from the nation's more highly regarded postsecondary institutions.

5.2 Difference-in-Difference

As previously noted, the OLS estimates presented in Table 2 most likely do not represent the true causal effect of a school district's decision to adopt performance on the selection of teachers into that district if districts that enact performance pay policies have underlying characteristics that are fundamentally different from districts that do not. In Table 3, we present the results of our difference-in-difference estimates which can better account for unobserved differences in districts that adopted teacher performance pay

[12] The mean SAT score is 1000 with a standard deviation of 200.

programs. By exploiting *change* in the same school districts' performance pay policies over time we uncover even stronger results than those obtained under the OLS analysis.

[Table 3 about here]

These difference-in-difference estimates reveal that new teachers selecting into performance pay districts graduated from colleges and universities with incoming freshman SAT scores around 30 points higher compared to the postsecondary credentials of teachers selecting into districts that did not adopt performance pay and based wages exclusively off of a single salary schedule. Note that the base sizes in Table 3 are smaller than those in Table 2 because the difference-in-difference estimate relies on a district being observed in both the 2003 and 2007 waves of SASS.

Next, we look at teachers with greater than four years of total teaching experience but fewer than four years of teaching experience in their current district. That is, we look at the extent to which performance pay may also attract experienced teachers with degrees from higher ranked institutions to make lateral moves into a performance pay district. In Table 4, we find that performance pay is also effective in attracting more experienced teachers who graduated from higher quality undergraduate institutions.

[Table 4 about here]

While the effect of performance pay is not as strong with these more experienced teachers, it is still significant. Compared to experienced teachers in non-performance pay districts, experienced teachers new to a performance pay district earned their baccalaureate degrees from institutions with SAT scores that are 17 points higher on average. Again, this result is robust to using either the 75th percentile or the 25th percentile of the postsecondary institution's average SAT score. The magnitude of this result is roughly half of the result for

truly new teachers (i.e. those with less than four years of total teaching experience). These results suggest that a school district which implements performance pay in order to increase the quality of its workforce, can expect to attract experienced teachers outside of the district with a stronger undergraduate profile in addition to increasing its appeal to freshly minted graduates seeking their first teaching position.

5.3 Robustness Checks

To obtain a representative teacher sample for the SASS, schools of all sizes across the country were sampled. The schools were then linked to the corresponding district in the SASS public school district survey. To ensure that our results are not being driven by a number of small districts, we undertake further sample restrict the sample used in our difference-in-difference estimation in three different ways in Table 5.

[Table 5 about here]

Column 1 of Table 5 is the baseline specification where we include districts of all sizes. In columns 2 and 3, we restrict the samples to those districts with at least 5 teacher observations and at least 10 teacher observations in the SASS respectively. Our findings are robust across these sample restrictions. Next, we subject our hypothesis that performance pay acts on teacher selection to a placebo test. Specifically, if performance pay were implemented by a district some time between 2003 and 2007, then we should not expect to find a change in the quality of the undergraduate institution for those teachers with greater than four years of experience in that same district since they would already have been employed by the district prior to the district's adoption of performance pay. The results of this placebo test are presented in Table 6 and they confirm our original hypothesis that teacher pay-for-performance has an independent effect on teacher selection behavior.

Across the SAT median, SAT 25th percentile, and SAT 75th percentile, we find no change in the quality of the undergraduate institution among teachers with more than four years of experience in the district.

[Table 6 about here]

Finally, we test the suitability of the identifying assumption for the DiD estimator – that is, that the trend in the quality of teacher recruitment for our treatment group of districts (those that adopted performance pay) would be the same as the control group that opted not to enact a performance pay program. To carry out this test, we restrict the sample to our control and treatment districts in our DiD specification in Table 3 to the 1999 and 2003 cycles of the SASS. We then estimate the following equation –

$$Y_{dst} = \beta_0 + 2003_{dst}\beta_1 + PerformancePay_{dst}\beta_2 + 2003*PerformancePay_{dst}\beta_3 + X_{dst}\beta_4 + \mu_s + \varepsilon_{dst} \quad (3)$$

Columns 1 and 3 in Table 7 show that for the SAT median and SAT 25th percentile, there is no statistically significant difference between the trends in the new teacher cohorts recruited by control and treatment districts between 1999 and 2003.

[Table 7 about here]

Column 2 of Table 7 shows that there may be a difference of twelve SAT points in the pre-treatment trends if teacher quality is measured by the incoming class' SAT 75th percentile of the teacher's undergraduate institution. Overall, the finding of this robustness check in trend assumptions suggests that prior to the implementation of performance pay between 2003 and 2007, performance pay districts (i.e. treatment districts) may have been recruiting slightly, but only slightly, higher quality teachers between 1999 and 2003. Nevertheless, the magnitude of this difference in the trends between treatment and control

districts is not nearly large enough to wash away the strength of our DiD estimates that districts implementing teacher pay-for-performance between 2003 and 2007 oversaw a notable increase in the quality of their new teacher hires.

6. Conclusion and Policy Implications

This paper is the first to document the relationship between the use of teacher performance pay and the quality of new teacher hires sorting into U.S. school districts. Drawing on two waves of SASS data we demonstrate that, on average, districts implementing performance pay to reward excellence in teaching secured new teachers who graduated from colleges and universities with average incoming SAT scores that were about 30 points higher than the new teacher cohorts hired by districts that did not introduce such performance pay programs. These findings are robust to a number of model specifications and persist even when we control for a district's use of other (non-performance-based) remuneration and pay incentive strategies such as: market-pay, hard-to-staff schools pay, and more generous baseline teacher pay.

Given the lack of attention paid by the scholarly literature to the relationship between pay reform and teacher recruitment and selection, these findings are an important first step toward assessing whether incentive pay can reshape candidate entry into the teaching profession. When one considers the decline in the academic aptitude of new teachers in the U.S. since about 1950 (Corcoran, Evans, and Schwab 2004a, 2004b; Hoxby and Leigh 2004) in the context of decades of studies demonstrating a link between teachers' academic aptitude and their classroom effectiveness (Clotfelter, Ladd, and, Vigdor 2007; Ehrenberg and Brewer 1994; Ferguson and Ladd 1996; Greenwood, Hedges, and,

Laine 1996; Hanushek 1981; Rockoff et al. 2011), our findings should encourage other researchers to pay greater attention to investigating whether and how compensation reform shapes districts' teacher recruitment efforts. For example, future studies might examine whether public education's use of fixed, versus variable pay, contributes to what some have labeled the "feminization" of the teaching profession in the U.S. (Hess 2010). A number of experimental studies have found that women are far more likely to prefer fixed versus variable pay schemes than are men (Dohmen and Falk 2011; Niederle and Vesterlund 2007) which raises an interesting question about whether public education's historic reliance on the single salary schedule dissuades talented male candidates from entering the profession. More importantly, future research could improve significantly upon this study by employing measures of teacher quality beyond academic aptitude/college selectivity, such as a teacher's actual value-added effectiveness. Chingos and West (2012) have already shown that more highly effective teachers who exit the profession typically enter into more lucrative post-teaching careers. Likewise, scholars might examine whether school districts that enact performance-based pay programs are better positioned to retain their most effective educators.[13]

Although we are confident that our results demonstrate that higher quality candidates do differentially prefer and ultimately sort into districts that offer opportunities to earn performance-based compensation, there are a number of important limitations to our study. First, the simple dichotomous measure available in the SASS survey differentiating a performance pay from a non-performance pay district preclude us from

[13] Economic theory would suggest that more productive employees prefer variable to fixed-pay systems. For a discussion of this theory specific to the teaching profession see Dohmen and Falk (2010).

drawing any firm conclusions about the optimal design (e.g. size, scope, reward structure) of a teacher pay-for-performance program whose chief aim is improving the quality of teachers recruited into a school district. Indeed, the precision of our estimates of the causal effect of a district adopting performance pay on the quality of new teacher hires is subject to considerable uncertainty since the SASS does not provide information on the size of the bonuses, whether they are group- or individual-based, and the fraction of teachers who routinely qualify for them.

Although these issues pose significant limitations to what we can learn from this current study, we do not believe that these concerns preclude policy-makers from drawing important conclusions about the relationship between teacher selection and performance pay reform. After all, in light of the fact that most performance pay initiatives in K-12 education today are small relative to a teacher's overall salary, the evidence provided here may actually understate the potential policy impact that a larger incentive-based pay system could have on selection into the teaching profession over the long term. Given that more highly-effective teachers have been shown to prefer individual performance pay systems (Perez, Muraki, and Loeb 2011) and that today's top college graduates in the U.S. say they will only consider a teaching career once performance bonuses up to 20 percent of salary are in place (Auguste, Kihn, and, Miller 2010), we suspect our empirical findings are estimating the mere floor that performance pay systems could have on teacher recruitment in the U.S.

References

Aaronson, Daniel, Lisa Barrow, and William Sander. 2007. Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics* 25(1): 99-135.

Anderson, Nick. 2010. Study undercuts teacher bonuses. *The Washington Post*. 22 September 2010: Met 2 Edition.

Atkinson, Adele, Simon Burgess, Bronwyn Croxson, Paul Gregg, Carol Propper, Helen Slater and, Deborah Wilson. 2009. Evaluating the impact of performance-related pay for teachers in England. *Labour Economics* 16(3): 251–61.

Auguste, Byron, Paul Kihn and, Matt Miller. 2010. Closing the talent gap: attracting and retaining top third graduates to a career in teaching. Washington, D.C.: McKinsey & Company.

Ballou, Dale. 2001. Pay for performance in public and private schools. *Economics of Education Review* 20(1): 51–61.

Ballou, Dale and Michael Podgursky. 1997. Teacher pay and teacher quality. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Berry, Barnett. 1988. Labor market choices and teacher reform: policy options for the public schools of the twenty-first century. *Teaching and Teacher Education* 4(1): 71-81.

Bowen, Daniel H., Buck, Stuart, Deck, Cary, Mills, Jonathan N. and, James V. Shuls. 2013. Risky business: an experimental evaluation of teacher risk preferences (EDRE Working Paper 2013-01).

Bowen, Daniel H. and Jonathan N. Mills. 2012. Does merit pay attract a different type of teacher? Evidence from a lab experiment. Working Paper.

Brewer, Dominic J., Eric R. Eide and, Ronald G. Ehrenberg. 1999. Does it pay to attend an elite private college? Cross-cohort evidence on the effects of college type on earnings. *Journal of Human Resources* 34(1): 104-23.

Buck, Stuart and Jay P. Greene. 2011. Blocked, diluted, and co-opted: Interest groups wage war against merit pay. *Education Next* 11(2): 27-32.

Chetty, Raj, John Friedman and, Jonah Rockoff. 2011. The long term impacts of teachers: Teacher value-added and student outcomes in adulthood. NBER Working Paper No. 17699.

Chingos, Matthew and Martin West. 2012. Do more effective teachers earn more outside of the classroom? *Education Finance and Policy* 7(1): 8–43

Clotfelter, Charles T., Ladd, Helen F. and, Jacob L. Vigdor. 2007. Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review* 26(6): 673-82.

Coleman, James S. 1966. Equality of Educational Opportunity Study. Report Commissioned by the United States Department of Health, Education, and Welfare.

Corcoran, Sean P., William N. Evans and, Robert Schwab. 2004a. Women, the labor market, and the declining relative quality of teachers. *Journal of Policy Analysis and Management* 23 (3): 449-70.

Corcoran, Sean P., William N. Evans and, Robert Schwab. 2004b. Changing labor-market opportunities for women and the quality of teachers, 1957-2000. *American Economic Review* 94(2): 230-35.

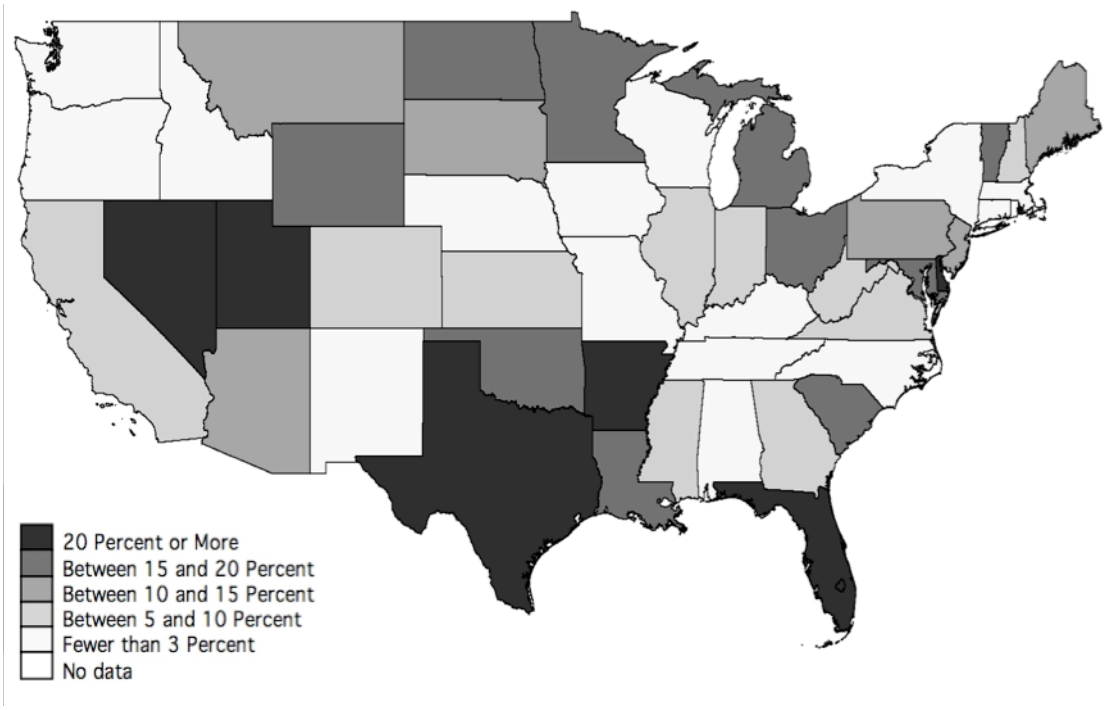
Dohmen, Thomas and Armin Falk. 2010. You get what you pay for: Incentives and selection in the education system. *Economic Journal* 120(546): F256-F271.
Dohmen, Thomas and Armin Falk. 2011. Performance pay and multidimensional sorting: Productivity, preferences, and gender. *American Economic Review* 101(2): 556-90.
Duncan, Arne. 2011. Working toward wow: A vision for a new teaching profession Speech delivered to the National Board of Professional Teaching Standards, July 29, 2011. Available online at: http://www.edweek.org/media/nbpts_ad_remarks-blog.pdf.
Ehrenberg, Ronald and Dominic Brewer. 1994. Do school and teacher characteristics matter? Evidence from highs school and beyond. *Economics of Education Review* 13(1): 1-17.
Eriksson Tor, Sabrina Teyssier and, Marie-Claire Villeval. 2009. Self-selection and the efficiency of tournaments. *Economic Inquiry* 47(3): 530-48.
Ferguson, Ronald F. and Helen F. Ladd. 1996. *How and why money matters: An analysis of Alabama schools*. Washington, D.C.: Brookings Institution Press.
Figlio, David N. 1997. Teacher salaries and teacher quality. *Economic Letters* 55(1): 267-71.
Figlio, David N. 2002. Can public schools buy better-qualified teachers? *Industrial Labor Relations Review* 55(4): 686-99.
Figlio, David N. and Lawrence W. Kenny. 2007. Individual teacher incentives and student performance. *Journal of Public Economics* 91(5-6): 901-14.
Fryer, Roland G. Forthcoming. Teacher incentives and student achievement: Evidence from New York City. *Journal of Labor Economics.*
Fryer, Roland G., Steven D. Levitt, John List, and, Sally Sadoff. 2013. Enhancing the efficacy of teacher incentives through loss aversion: A field experiment. NBER Working Paper No. 18237.
Goldhaber, Dan. 2002. The mystery of good teaching. *Education Next* 2(1): 50-55.
Goldhaber, Dan and Dominic Brewer. 1997. Why don't schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity. *Journal of Human Resources* 32(3): 505-23.
Goldhaber, Dan and Dominic Brewer. 2000. Does teacher certification matter? High school teacher certification Status and student achievement. *Educational Evaluation and Policy Analysis* 22(2): 129-45.
Goldhaber, Dan, Daniel Player, Michael DeArmond, and, Hyung-Jai Choi. 2008. Why do so few public school districts use merit pay? *Journal of Education Finance* 33(3): 262-289.
Goldhaber, Dan and Jane Hannaway. 2009. Editors. Creating a new teaching profession. Washington, DC: Urban Institute Press.
Goldhaber, Dan and Joe Walch. 2012. Strategic pay reform: A student outcomes-based evaluation of Denver's ProComp teacher pay initiative. *Economics of Education Review* 31(6): 1067–83.
Greenwood, Rob, Larry Hedges and, Richard Laine. 1996. The effect of school resources on student achievement. *Review of Educational Research* 66(3): 361-96.
Glewwe, Paul, Nauman Ilias, and, Michael Kremer. 2010. Teacher Incentives. *American Economic Journal: Applied Economics* 2(3): 205–27.
Hanushek, Eric A. 1981. Throwing money at schools. *Journal of Policy Analysis and Management* 1(1): 19-41.
Hanushek, Eric A. and Steven G. Rivkin. 2004. How to improve the supply of high quality

teachers. In Diane Ravitch, ed., Brookings papers on education policy, 7-25. Washington, DC: Brookings Institution Press.
Hendricks, Matthew. 2011. Performance pay and teacher selection: Do performance pay programs attract higher-ability teachers? Unpublished doctoral dissertation (University of Minnesota).
Hess, Frederick M. 2004. Common sense school reform. New York, NY: Palgrave MacMillan.
Hess, Frederick M. 2010. The same thing over and over: How school reformers get stuck in yesterday's ideas. Harvard University Press.
Hoxby, Caroline and Andrew Leigh. 2004. Pulled away or pushed out? Explaining the decline of teacher aptitude in the U.S. *American Economic Review* 94(2): 236-40.
Imberman, Scott A. and Michael F. Lovenheim. 2012. Incentive strength and teacher productivity: Evidence from a group-based teacher incentive pay system. NBER Working Paper No. 18439.
Koedel, Cory and Julian R. Betts. 2011. Does student sorting invalidate value-added models of teacher effectiveness? An extended analysis of the Rothstein critique. *Education Finance and Policy* 6(1): 18-42.
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): 1346-61.
Lazear, Edward P. 2003. Teacher Incentives. *Swedish Economic Policy Review* 10, 197–213.
Leigh, Andrew. 2013. The economics and politics of teacher merit pay. *CESifo Economic Studies* 59(1): 1-31.
Leuven, Edwin, Hessel Oosterbeek, Joep Sonnemans and, Bas van der Klaauw. 2011. Incentives versus sorting in tournaments: Evidence from a field experiment. *Journal of Labor Economics* 29(3): 637-58.
McCaffrey, Daniel F., Daniel Koretz, J. R. Lockwood and, Laura S. Hamilton. 2006. Evaluating value-added models for teacher accountability. Washington, D.C.: RAND Corporation.
Muralidharan, Karthik and Venkatesh Sundararaman. 2011. Teacher performance pay: Experimental evidence from India. *Journal of Political Economy* 119(1): 39–77.
Niederle, Muriel and Lise Vesterlund. 2007. Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics* 122(3): 1067-1101.
Perez, Maria, Mari Muraki and, Susanna Loeb. 2011. Do school and teacher value-added measures correlate with their performance-based pay preferences. Working Paper.
Player, Daniel. 2009. Monetary returns to academic ability in the public teacher labor market. *Economics of Education Review* 28(2): 277-85.
Podgursky, Michael J. and Matthew Springer. 2007. Teacher performance pay: A review. *Journal of Policy Analysis and Management* 26(4): 909-49.
Ravitch, Diane. 1974. The great school wars: New York City, 1805-1973. Baltimore: Johns Hopkins University Press.
Rivkin, Steven G., Eric A. Hanushek and, John F. Kain. 2005. Teachers, schools, and academic achievement. *Econometrica* 73(2): 417-58.
Rockoff, Jonah, Brian A. Jacob, Thomas J. Kane and, Douglas O. Staiger. 2011. Can you

recognize an effective teacher when you recruit one? *Education Finance and Policy* 6(1): 43-74.
Rothstein, Jesse. 2009. Student sorting and bias in value added estimation: selection on observables and unobservables. *Education Finance and Policy* 4(4): 537-71.
Rothstein, Jesse. 2010. Teacher quality in educational production: Tracking, decay, and student achievement. *Quarterly Journal of Economics* 125(1): 175-214.
Sanders, William and June C. Rivers. 1996. Cumulative and residual effects of teachers on future student academic achievement. Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center.
Sawchuck, Stephen. 2010. Merit pay found to have little effect on achievement. *Education Week* (9 September 2010).
Springer, Matthew G., Dale Ballou, Laura S. Hamilton, Vi-Nhuan Le, J. R. Lockwood, Daniel F. McCaffrey, Matthew Pepper and, Brian M. Stecher. 2010. Teacher pay for performance: Experimental evidence from the project on incentives in teaching. Nashville, TN: National Center on Performance Incentives at Vanderbilt University.
West, Kristine Lamm and Elton Mykerezi. Teachers' unions and compensation: The impact of collective bargaining on salary schedules and performance pay schemes. *Economics of Education Review* 30(1): 99-108.
Winters Marcus A., Gary Ritter, Ryan Marsh, Jay P. Greene and, Marc Holley. 2008. The impact of performance pay for public school teachers: Theory and evidence. Program on Education Policy and Governance Working Papers Series, Harvard University.
Woessmann, Ludger. 2011. Cross-country evidence on teacher performance pay. *Economics of Education Review* 30(3): 404–18.

Appendix A: Tables & figures

Figure 1: Percentage of School Districts Using Performance Pay: by State, 2007-2008



Source: National Center for Education Statistics, Schools and Staffing Survey, 2007-2008

Table 1: 2007 SASS, School District Summary Statistics

	Performance Pay District	Non Performance Pay District	Difference
Union	0.39	0.59	-0.20***
Urban	0.43	0.31	0.12**
Graduation Rate	0.68	0.74	-0.06**
Expenditures per Student	$11,607	$11,581	$26
Free Lunch Student %	0.44	0.39	0.05***
Student Teacher Ratio	14.92	13.94	0.99***
High Maintenance District	0.18	0.24	-0.07
Low Maintenance District	0.36	0.26	0.10**
Overachieving District	0.38	0.39	-0.01
Underachieving District	0.08	0.10	-0.02
Male Teacher %	0.29	0.32	-0.03
Age of Teacher	31.55	29.92	1.63**
White Teacher %	0.90	0.96	-0.06***
Teachers with Master's Degree %	0.32	0.22	0.09
New Teacher Salary	$33,239	$33,418	-$179
District Offers Signing Bonus	0.22	0.09	0.13***
District Offers Relocation Assistance	0.09	0.05	0.04***
District Offers Loan Forgiveness	0.09	0.02	0.07***
District Rewards Hard-to-staff	0.33	0.05	0.28***
District Rewards Field Shortages	0.49	0.19	0.30***
Observations	140	1520	1660

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Data is weighted using SASS provided weights

Note: Sample sizes rounded to nearest 10 for NCES confidentiality purposes

Table 2: OLS Estimates of Performance Pay on Quality of Teachers w/ < 4 Years of Total Experience

	SAT Median	SAT 75th Percentile	SAT 25th Percentile
Performance Pay District	16.963**	18.832**	15.094*
	(7.898)	(8.366)	(7.855)
Free Lunch Percentage	-81.765***	-76.916***	-86.614***
	(13.111)	(13.363)	(13.153)
Expenditures Per Student	1.790*	1.626*	1.955**
	(0.926)	(0.965)	(0.908)
Graduation Rate	-41.107*	-35.821	-46.393**
	(21.384)	(22.323)	(20.983)
Student-Teacher Ratio	-0.176	-0.439	0.086
	(0.643)	(0.668)	(0.626)
Teacher FTE	0.000	0.000	0.001*
	(0.000)	(0.000)	(0.000)
Union	25.587***	26.269***	24.904***
	(5.780)	(6.328)	(5.476)
Urban	-7.098	-5.097	-9.100*
	(5.066)	(5.264)	(5.042)
White Teacher %	12.249	11.054	13.444
	(13.674)	(15.436)	(13.597)
Male Teacher %	10.775	8.675	12.874
	(10.491)	(10.312)	(10.962)
Age of Teacher	-1.626***	-1.853***	-1.398**
	(0.524)	(0.534)	(0.531)
Teachers with Master's %	19.010**	15.679	22.341**
	(8.939)	(9.599)	(8.705)
New Teacher Salary	0.003***	0.004***	0.003***
	(0.001)	(0.001)	(0.001)
Observations	1660	1660	1660
R-Square	0.231	0.210	0.248

Also includes state fixed effects, other (non performance-based) pay recruitment incentive dummies, and district characteristic (e.g. underachieving, overachieving, and low maintenance) dummies

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Data is weighted using SASS provided weights

Note: Sample sizes rounded to nearest 10 for NCES confidentiality purposes, using 2007 SASS data

Table 3: DiD Estimates of Performance Pay on Quality of Teachers w/ < 4 Years of Total Experience

	(1) SAT Median	(2) SAT 75th Percentile	(3) SAT 25th Percentile
2007*Performance Pay District	32.263**	30.964***	33.562**
	(12.872)	(11.586)	(14.650)
2007	-9.364	-10.673*	-8.055
	(6.542)	(6.314)	(7.155)
Performance Pay District	-16.362**	-13.567*	-19.157**
	(8.041)	(6.920)	(9.594)
Free Lunch Percentage	-49.181***	-45.751***	-52.612***
	(17.701)	(17.092)	(18.995)
Expenditures Per Student	2.070**	1.695*	2.445**
	(1.020)	(0.939)	(1.135)
Graduation Rate	28.269	24.823	31.716
	(24.437)	(23.213)	(26.866)
Student-Teacher Ratio	0.102	0.160	0.044
	(0.185)	(0.180)	(0.193)
Teacher FTE	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)
Union	22.470***	19.597***	25.343***
	(6.526)	(6.792)	(6.403)
Urban	8.758*	8.868*	8.649*
	(4.945)	(5.098)	(5.021)
White Teacher %	35.224**	43.988***	26.460*
	(15.562)	(16.260)	(15.468)
Male Teacher %	7.443	2.735	12.152
	(12.336)	(12.192)	(13.213)
Age of Teacher	-0.554	-0.942*	-0.166
	(0.529)	(0.525)	(0.590)
Teachers with Master's %	10.293	11.059	9.527
	(9.412)	(9.315)	(10.518)
New Teacher Salary	0.004***	0.004***	0.005***
	(0.001)	(0.001)	(0.001)
Observations	1130	1130	1130
R^2	0.244	0.237	0.244

Also includes state fixed effects, other (non performance-based) pay recruitment incentive dummies, and district characteristic (e.g. underachieving, overachieving, and low maintenance) dummies

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Data is weighted using SASS provided weights

Note: Sample sizes rounded to nearest 10 for NCES confidentiality purposes

Table 4: DiD Estimates of Performance Pay on Quality of Teachers with >= 4 Years of Total Experience and < 4 Years of Experience in their Current District

	(1) SAT Median	(2) SAT 75th Percentile	(3) SAT 25th Percentile
2007*Performance Pay District	17.726*	16.998*	18.453*
	(9.504)	(9.828)	(10.370)
2007	10.110	7.327	12.894**
	(6.194)	(6.104)	(6.477)
Performance Pay District	-18.783**	-19.147**	-18.419**
	(7.288)	(8.153)	(7.309)
Free Lunch Percentage	-57.575***	-55.747***	-59.404***
	(13.722)	(13.921)	(13.913)
Expenditures Per Student	1.363	1.489	1.237
	(1.019)	(0.986)	(1.075)
Graduation Rate	-10.276	-8.449	-12.104
	(24.847)	(24.264)	(25.938)
Student-Teacher Ratio	0.226**	0.283***	0.169*
	(0.096)	(0.095)	(0.100)
Teacher FTE	0.000	0.001	-0.000
	(0.000)	(0.000)	(0.000)
Union	13.911**	13.553**	14.268**
	(5.718)	(5.885)	(6.176)
Urban	8.517**	7.577*	9.457**
	(4.099)	(4.238)	(4.146)
White Teacher %	10.265	16.528	4.003
	(13.086)	(13.807)	(12.708)
Male Teacher %	-6.426	-6.335	-6.517
	(7.836)	(8.230)	(7.764)
Age of Teacher	-0.185	-0.356	-0.013
	(0.396)	(0.399)	(0.410)
Teachers with Master's %	41.262***	42.154***	40.370***
	(8.858)	(9.198)	(8.798)
New Teacher Salary	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)
Observations	1130	1130	1130
R^2	0.244	0.237	0.244

Also includes state fixed effects, other (non performance-based) pay recruitment incentive dummies, and district characteristic (e.g. underachieving, overachieving, and low maintenance) dummies

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Data is weighted using SASS provided weights

Note: Sample sizes rounded to nearest 10 for NCES confidentiality purposes

Table 5: Robustness Check for DiD Estimates of Performance Pay on Quality of Teachers

	(1) SAT Median - Districts w/ at least 1 Teacher Observation	(2) SAT Median - Districts w/ at least 5 Teacher Observations	(3) SAT Median - Districts w/ at least 10 Teacher Observations
2007*Performance Pay District	32.263**	27.797***	29.901**
	(12.872)	(9.659)	(13.230)
2007	-9.364	-10.888*	-13.144**
	(6.542)	(5.675)	(5.257)
Performance Pay District	-16.362**	-23.386***	-15.375**
	(8.041)	(6.378)	(7.165)
Free Lunch Percentage	-49.181***	-48.990***	-52.440***
	(17.701)	(14.199)	(13.905)
Expenditures Per Student	2.070**	2.660***	3.265***
	(1.020)	(0.917)	(0.678)
Graduation Rate	28.269	21.747	12.848
	(24.437)	(22.111)	(16.263)
Student-Teacher Ratio	0.102	-0.130	-0.498***
	(0.185)	(0.174)	(0.060)
Teacher FTE	-0.000	0.001*	0.001***
	(0.000)	(0.000)	(0.000)
Union	22.470***	23.379***	12.754***
	(6.526)	(5.703)	(4.134)
Urban	8.758*	15.141***	5.953
	(4.945)	(3.739)	(3.584)
White Teacher %	35.224**	71.426***	77.065***
	(15.562)	(14.188)	(14.575)
Male Teacher %	7.443	-9.366	29.981**
	(12.336)	(12.439)	(13.695)
Age of Teacher	-0.554	-0.219	1.301**
	(0.529)	(0.648)	(0.599)
Teachers with Master's %	10.293	12.400	-4.237
	(9.412)	(11.425)	(14.601)
New Teacher Salary	0.004***	0.004***	0.003***
	(0.001)	(0.001)	(0.001)
Observations	1130	940	500
R^2	0.244	0.250	0.338

Also includes state fixed effects, other (non performance-based) pay recruitment incentive dummies, and district characteristic (e.g. underachieving, overachieving, and low maintenance) dummies

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Data is weighted using SASS provided weights

Note: Sample sizes rounded to nearest 10 for NCES confidentiality purposes

Table 6: Robustness Check for DiD Estimates of Performance Pay on Quality of Teachers with >= 4 Years of Total Experience and >= 4 Years of Experience in their Current District

	(1) SAT Median	(2) SAT 75th Percentile	(3) SAT 25th Percentile
2007*Performance Pay District	4.372	-3.980	12.724
	(10.856)	(10.455)	(15.521)
2007	-2.650	-3.109	-2.190
	(5.075)	(4.934)	(5.353)
Performance Pay District	-15.258**	-7.321	-23.196*
	(7.522)	(7.034)	(13.429)
Free Lunch Percentage	-32.838***	-32.755***	-32.922***
	(11.362)	(11.261)	(11.716)
Expenditures Per Student	2.305***	2.281***	2.329**
	(0.872)	(0.811)	(0.948)
Graduation Rate	29.253*	27.086	31.419*
	(16.673)	(16.558)	(17.426)
Student-Teacher Ratio	-0.326***	-0.350***	-0.301***
	(0.088)	(0.092)	(0.087)
Teacher FTE	0.000	0.001*	0.000
	(0.000)	(0.000)	(0.000)
Union	18.860***	19.762***	17.957***
	(3.418)	(3.365)	(3.643)
Urban	14.115***	14.109***	14.122***
	(3.680)	(3.730)	(3.777)
White Teacher %	71.304***	80.924***	61.684***
	(10.540)	(10.627)	(10.666)
Male Teacher %	-5.551	-7.172	-3.929
	(6.127)	(6.206)	(6.326)
Age of Teacher	-0.994***	-1.108***	-0.880***
	(0.223)	(0.232)	(0.227)
Teachers with Master's %	8.179	10.800**	5.558
	(5.365)	(5.424)	(5.586)
New Teacher Salary	0.002***	0.002***	0.002**
	(0.001)	(0.001)	(0.001)
Observations	1780	1780	1780
R^2	0.285	0.289	0.286

Also includes state fixed effects, other (non performance-based) pay recruitment incentive dummies, and district characteristic (e.g. underachieving, overachieving, and low maintenance) dummies

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Data is weighted using SASS provided weights

Note: Sample sizes rounded to nearest 10 for NCES confidentiality purposes

Table 7: Trends Assumption Robustness Check for DiD Estimates of Performance Pay on Quality of Teachers using 1999 and 2003 Districts

	(1) SAT Median	(2) SAT 75th Percentile	(3) SAT 25th Percentile
2007*Performance Pay District	9.894	12.436*	7.352
	(6.194)	(6.777)	(5.840)
2007	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Performance Pay District	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Free Lunch Percentage	-31.847**	-30.027*	-33.666**
	(15.481)	(15.721)	(15.855)
Expenditures Per Student	0.265	-0.365	0.895
	(1.436)	(1.399)	(1.542)
Graduation Rate	55.740*	53.222*	58.258*
	(28.438)	(27.275)	(31.124)
Student-Teacher Ratio	0.350**	0.352**	0.348**
	(0.167)	(0.164)	(0.171)
Teacher FTE	0.001*	0.001**	0.001
	(0.000)	(0.000)	(0.000)
Union	5.463	5.097	5.829
	(6.807)	(6.352)	(7.493)
Urban	11.769**	9.860	13.677**
	(5.909)	(6.243)	(5.883)
White Teacher %	25.368	34.857	15.879
	(20.208)	(22.055)	(18.919)
Male Teacher %	-26.435	-32.550*	-20.319
	(18.229)	(18.052)	(19.351)
Age of Teacher	1.405**	1.254**	1.556**
	(0.647)	(0.600)	(0.723)
Teachers with Master's %	11.603	8.507	14.700
	(14.664)	(15.110)	(15.158)
New Teacher Salary	0.002	0.002	0.002
	(0.002)	(0.002)	(0.002)
Observations	780	780	780
R^2	0.242	0.251	0.241

Also includes state fixed effects, other (non performance-based) pay recruitment incentive dummies, and district characteristic (e.g. underachieving, overachieving, and low maintenance) dummies
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Data is weighted using SASS provided weights
Note: Sample sizes rounded to nearest 10 for NCES confidentiality purposes

Appendix B: Data and Replication Materials

The dataset utilized by this study is confidential as required by the Institute for Educational Sciences' (IES) restricted use data policy, which expressly precludes the authors from sharing details from the Schools and Staffing Surveys (SASS). Readers who wish to replicate the analysis will need to apply for and obtain a restricted-use license directly from IES: IES Data Security Office, U.S. Department of Education/NCES, 1990 K. Street NW, Room 9060 Washington, DC 20006; 202-502-7307, IESData.Security@ed.gov.

A data appendix with results and copies of the computer programs used to generate the results presented in the manuscript are readily available. Interested parties may contact:

Michael Jones
Research Assistant Professor
Department of Economics
323 Lindner Hall
University of Cincinnati
Cincinnati, OH 45221
513-276-7093 (office)
m.jones@uc.edu