The Effect of Teacher Performance Pay on Adult Outcomes in the United States

Timothy N. Bond
Purdue University

Kevin J. Mumford
Purdue University

October 2016

Abstract

We estimate the effect of exposure to teacher pay-for-performance programs on adult outcomes. We construct a comprehensive data set of schools which have implemented teacher performance pay programs across the United States since 1986, and use our data to calculate the fraction of students in each grade in each state who are affected by a teacher performance pay program in a given year. We then calculate the expected years of exposure for each birth state-grade cohort in the American Community Survey. Cohorts with more exposure earn lower wages as adults. This negative effect is concentrated on women and high school graduates with no significant effect for men. We find possible positive effects for high school drop outs.

JEL Codes: I21, J24

We are grateful to participants at the 2015 Association for Public Policy and Management Fall Research Conference, the 2016 European Association of Labor Economists Conference, and seminar participants at Brigham Young University, Purdue University, and the University of Nottingham for their helpful comments. We would like to thank Mary Kate Bastitch, Jacklyn Buhrmann, Kendall Kennedy, Sophie Shen, and Wanling Zou for excellent research assistance. Any remaining errors are our own.

I Introduction

Approximately 100 percent of public school teachers work for a school district with a salary schedule (Podgursky 2007). These salary schedules differentiate pay among teachers by experience, seniority, and credentials, but not generally by observed performance. Education reformers have long viewed this as problematic for two reasons. First, the classroom environment presents a classic case of moral hazard: it is difficult for a principal to observe the collective set of actions taken by a teacher over the school year and to know what the optimal set of actions would have been (Neal 2011). Second, the characteristics that do differentiate pay under the salary schedule have little correlation with teacher performance (Hanushek 2003).

As a consequence, and in conjunction with the advent of modern standardized testing, policy makers have increasingly sought to tie teacher pay to student test performance.[1] Comprehensive reviews of teacher performance pay by Podgursky and Springer (2007) and Neal (2011) have found that teachers do respond to performance pay incentives. When teachers receive bonuses tied to a specific test, their students' scores on that test generally increase.[2] But, there are three major reasons to be skeptical of such analyses: (1) test scores can be manipulated by teaching to the test or by orchestrated cheating, (2) test scores do not reflect many important skills that can be taught, and (3) test scores are measured on ordinal scales which makes over-time or across-group comparisons unconvincing.

In this paper, we avoid these concerns by estimating the effect of teacher performance pay programs on wages and other adult outcomes. In contrast with the generally positive test score results, we find that one year of exposure to a teacher performance pay program in the United States leads to a decrease in adult wages of 0.9 percent on average. This estimate

[1] Tying teacher pay to test scores is only the latest initiative in the long-standing effort to differentiate teacher pay by ability. The previous iteration came through the movement towards "career ladders" in the 1980s, which gave higher salaries to teachers who obtained more education, outside credentials, and passed evaluations through classroom observations. These programs were generally deemed overly expensive, with little evidence of success on student outcomes, and were almost universally abolished by the early 1990's (Cornett and Gaines 1994).

[2] Although, it is important to note that this is not a universal finding. See the discussion in footnote 3

is obtained from a specification in which we control for per pupil spending, which means that adult wages are estimated to be 0.9 percent lower than would have occurred had the performance pay funds been given to the school to spend on other educational priorities. We do, however, find evidence that not all programs were harmful. Estimates from a random slope model suggest that 9 percent of states experienced positive wage effects from their programs conditional on per pupil spending.

The literature has focused on estimating the effect on student test scores for specific (generally state-level) teacher performance pay programs.[3] We have already discussed the shortcomings of this approach and there is strong empirical evidence to suggest these concerns are not trivial. Jacob and Levitt (2003) found evidence that teachers responded to the introduction of high-stakes testing in Chicago by changing student answers, i.e. cheating. Koretz and Barron (1998) found that the introduction of test-based accountability led to increases in student performance on the test that was used for school evaluations, but not on tests that were not. Even evaluating performance pay programs with tests that were not used for bonus schemes can be problematic. Teachers may direct resources to test-taking skills which have positive effects on all test scores, without any human capital gains (Neal 2013). Finally, Bond and Lang (2013) show that estimated effects in the black-white test score gap literature are not robust to order-preserving scale transformations of test scores. This implies that the test scale may itself be responsible for some of the differences in estimated effects across performance pay programs.

Few researchers have attempted to evaluate school policy using adult outcomes rather than test scores. Chetty, Friedman, and Rockoff (2014) examine the relationship between teacher quality and adult wages using teacher value-added. Heckman, Pinto, and Savelyev

[3]For example, Cooper and Cohn (1997) find a positive effect on test scores in South Carolina; Ladd (1999) finds a positive effect on test scores in Dallas; Dee and Keys (2004) find a positive effect on test scores in Tennessee; Winters, Greene, Ritter, and Marsh (2008) find a positive effect on test scores in Little Rock; Vigdor (2009) finds a positive effect on test scores in North Carolina; Sojourner, Mykerezi, and West (2014) find a positive effect on test scores in Minnesota; and Imberman and Lovenheim (2015) find a positive effect in Houston. Other studies have found no effect on test scores including Eberts, Hollenbeck, and Stone (2002) in Michigan, Lincove (2012) in Texas, and Fryer (2013) and Goodman and Turner (2013) in New York City.

(2013) use adult outcomes to evaluate the Perry Preschool program. Dobbie and Fryer (2016) use adult outcomes to evaluate the performance of charter schools. Card and Krueger (1998) find weak evidence that school resources are positively correlated with adult outcomes, while Jackson, Johnson, and Persico (2016) estimate that a 10 percent increase in per pupil spending causes a 7 percent increase in adult wages. In a closely related work, Lavy (2015) evaluates the effect of a high school teacher performance pay program experiment in Israel and finds that treated students have 7 percent higher earnings in their late 20's and early 30's than the control students. However, the Israeli program is not representative of the design of teacher performance pay programs in the United States. One important difference is the teacher performance pay in Israel was based on a student performance on a college-entrance exam and better performance on the test increased the likelihood that the treated students attend and graduate from college. It is possible that the entire effect in Lavy (2015) is simply a return to increasing college education, not a return to increased effort by high school teachers.

We construct a comprehensive school-level data, using multiple sources, of every teacher performance pay program in the United States from 1986-2007. To our knowledge, no other systematic compilation of teacher performance pay programs exists.[4] Rather than use the actual exposure to teacher performance pay programs for each individual (which we do not observe), we calculate the likely years of exposure by state and cohort. A benefit is that this alleviates the concern that schools that select into teacher performance pay programs may be better schools, or that states target teacher performance pay programs towards worse schools. We calculate the fraction of public school students in each state and year that are exposed to teacher performance pay. We then total this exposure for grades 1 through 12 for each birth-state cohort. The American Community Survey (ACS) provides us with the adult outcomes for these cohorts. We find that cohorts with a larger share of individuals enrolled in schools with performance pay programs earn lower wages on average as adults. We also

[4] There is some tracking of performance pay programs in the United States by Mathematica and the National Center on Performance Incentives, but these efforts are incomplete.

find evidence of important distributional effects with the negative wage effects concentrated among women. We find some evidence of positive effects for high school dropouts with the largest negative effects for those with a college degree, suggesting that teachers respond to performance pay incentives by redirecting attention from those who are easy to teach to those who are more difficult.

The remainder of this article is organized as follows. Section II presents a model of teacher attention that suggests that teachers shift attention towards low-ability students in response to a performance pay program. Section III describes our two-year data collection effort which resulted in the most comprehensive dataset of teacher performance pay programs ever constructed. We then discuss the trends in teacher performance pay in the United States over the last three decades. Section IV outlines our estimation strategy, Section V presents our results, and Section VI presents our conclusions.

II Model of Teacher Attention

Theoretical analysis of teacher incentive pay plans have focused on the "multi-tasking problem." In short, teachers can respond to the introduction of a high-stakes test in two ways. The first is by increasing effort towards teaching general knowledge and skills that are both measured by the test and valued by the market. The second is to spend time "coaching" students on techniques that are valuable only in increasing scores on the specific assessment device, and not generally valued on the market.[5] As Neal (2011) shows, the theoretical impact of introducing pay-for-performance on student outcomes is ambiguous. When the test is difficult to coach, the increase in total effort dominates the diversion of teacher resources towards coaching. When it is not, teachers decrease their time on human capital building activities and replace them with coaching, which can lower labor market outcomes. Thus

[5] For example, it was well known that the early 2000s version of the verbal section of the Graduate Record Exam (GRE), used by many graduate schools as part of the admissions decision, relied heavily on English words with Latin roots. Coaching would involve teaching only Latin-root words despite the likely equal labor market value of words of Germanic origin.

the actual effect of pay-for-performance on student welfare is an empirical question which we seek to answer.

Here we show a second potential outcome of teacher pay-for-performance. Introducing such a policy can have important distributional effects on teacher effort. Abstracting from the impact of pay-for-performance on the total provision of effort and its distribution on tasks, suppose that a teacher has one unit of time that she must allocate between a student who is easy to teach (E) and a student who is difficult (D). The student may be easy because he is of innately higher ability, has fewer behavioral problems, or shares a similar background, such as race or gender, with the teacher. Time spent on teaching the difficult student is more costly to the teacher.

A student who receive t units of teacher attention obtains human capital $h(t)$ with $h'(t) > 0$ and $h''(t) < 0$. The teacher is intrinsically motivated to produce human capital, valuing each unit produced at a regardless of the student who receives it. The teacher's utility function is

$$U = a[h(t_E) + h(1 - t_E)] - t_E - \lambda(1 - t_E) \tag{1}$$

where t_E is the time the teacher devotes to the easy student and $\lambda > 1$ reflects that time spent with the difficult student is more costly. It is straightforward to solve the teacher's optimization problem using a first order condition,

$$h'(t_E^*) - h'(1 - t_E^*) = \frac{1}{a}(1 - \lambda) \tag{2}$$

Since the right-hand side is negative, and $h(t)$ is concave, we know that the $t_E^* > \frac{1}{2}$. The teacher devotes more attention towards the easy student simply because they are easier to teach.

Now suppose we introduce a pay-for-performance plan that is tied to a *perfect* test. The education board simply observes h for both students and pays the teacher a bonus b for each unit of human capital she produced. Again, the board does not care which student earns h

in this example; it is simply rewarding the teacher for the total human capital she produces. The teacher's utility function becomes

$$U = (a + b)[h(t_E) + h(1 - t_E)] - t_E - \lambda(1 - t_E) \tag{3}$$

The solution to the maximization problem is now,

$$h'(t_E^*) - h'(1 - t_E^*) = \frac{1}{a + b}(1 - \lambda) \tag{4}$$

Compared to the solution without the bonus payments, the right-hand side is smaller in absolute value. Thus, t_E^* decreases. The teacher spends more time with difficult students.

The intuition behind this result can be thought of in terms of the labor/leisure trade-off for teachers. Teaching easy students is relatively more "leisure-ful" than teaching difficult students. Thus, given the curvature of $h(t)$, at the margin an additional ϵ unit of time invested in the difficult students would produce more human capital than were it to be invested in the easy students. Adding the bonus increases the weight a teacher places on human capital relative to her leisure. Since the returns are higher at the margin for difficult students in the previous equilibrium, performance pay induces the teacher to spend relatively more time with them.

The simple model demonstrates the impact on teacher's attention under a pseudo-value added framework. Many performance pay programs reward teachers for the fraction of students above a specific "proficiency" score, rather than just for simple gains in test scores. Under such a program, the distributional effects are even more straightforward. Teachers receive monetary benefits from raising marginal students about the testing threshold, and no benefits to improving students who are already well above it.[6]

While a given teacher performance pay program may increase or decrease the labor

[6]Depending on the structure of awards, teachers may also face an incentive to ignore the absolute worst students who have no hope of reaching the threshold.

market outcomes of students on average relative to the counter factual with the same increase in per pupil spending, we would expect that the students who are most difficult to teach would have relatively more positive effects than those who are easier to teach. This could include low ability students who have lower educational outcomes. Previous research has also suggested that teachers may view boys as more difficult to teach than girls due to differences in behavior (Cornwell, Mustard, and Parys 2013) or non-cognitive skills (Jacob 2002), and that students perform better when being taught by a teacher of the same race and gender (Dee 2004, Dee 2005). Given these factors, and that teachers are disproportionately white and female, we should expect to see more positive (less negative) effects on men and racial minorities.

III Data

One of the primary barriers to testing the impact of teacher incentive pay is the lack of data on its incidence. We address this problem by constructing, by hand using multiple sources, a panel dataset of schools whose teachers were paid for student test performance. We began with the (formerly) publicly available data from the Center for Education Compensation Reform. This provided us an incomplete list of programs in the United States. We supplemented this through searches of publicly available documents using Google and Lexus-Nexus, and a search of news articles through ProQuest Newstand. The ProQuest search was especially helpful in identifying small now-defunct performance pay programs. In instances where our search results do not identify the complete list of schools affected by a particular program, we contacted the appropriate government officials and in some instances submitted a Freedom of Information request.[7] As a consequence, for nearly every program in our database, we know the exact schools which were affected by performance pay pro-

[7]We initially contacted school superintendents as primary sources. While a few were very helpful, the vast majority said that did not have knowledge of programs that ended before they took their position. In some instances, teacher union leaders had excellent information about which schools were affected in each year.

grams.[8] This effort identified 7 state-wide programs, 138 district-wide programs, and 2,925 school-specific programs that were implemented in at least one academic year between 1986 and 2007.[9]

Unfortunately, we are able to observe the size of teacher bonuses for less than 20 percent of the programs we identified, and this was heavily skewed towards recent programs. It appears to us that the details of many earlier discontinued programs have been lost to time. We were more successful in identifying if the teacher bonus payments were tied to individual, group, or district-level performance. We found that nearly every program had some component of the bonus that was based on individual performance. Thus we believe it would be impossible to disentangle which type of incentives were important.

To get counts of students who were affected by each program, we merge our data with school enrollment counts from the National Center for Education Statistics Common Core of Data Universe Surveys (CCD). The CCD provides characteristics on each school in the United States in each year since the 1986-1987 school year. These characteristics include counts by grade, socioeconomic status, gender, and race. The CCD has 147,618 schools that enrolled students between 1986 and 2007.[10] Teachers at 25,680 of these schools received pay that was a result of student test score performance. The enrollment data allows us to calculate the number of students exposed to a teacher performance pay program. We deal with missing data by imputing grade-level enrollments from total school enrollments and grades offered at each school, where enrollment is distributed evenly over each grade offered.

In Figure 1, we plot the fraction of non-charter public school children in the United States who were enrolled in a school which offered performance pay from 1986-2007. Starting in the early 1990s, teacher performance pay programs have been adopted by an increasing number of schools. The figure shows that on average an additional 1 percentage point of

[8]We lack school-specific information for some programs in California, New York, Georgia, Louisiana, South Carolina, and Chicago. For these programs we instead use the fraction of schools affected in each year, and assign this value to each school in the state or district.

[9]For the above figures, each school is counted as being in a separate program if only a subset of schools within the relevant school district are selected into a district-wide program

[10]Note that there are only about 100,000 schools with enrolled students in any given year.

students are in schools with performance pay programs in each year since 1992. As of the 2007-2008 school year, 17 percent of students in the United States were being educated in such a manner.

Figure 2 maps the fraction of non-charter public school children who were enrolled in a school which offered performance pay by state in selected years. In the 1980s, only a small number of school districts in Arizona offered test-based teacher incentive pay. We see an increase in performance pay in the 1990s in part due to the adoption of large-scale programs in North Carolina, Florida, Georgia, and California, as well as several district-wide programs. The trend continues into the 2000s with the growth of programs, for example, in Arizona and Minnesota, and as a possible consequence of No Child Left Behind legislation making tests more prevalent on which teacher performance pay can be based.[11]

Adult Outcome Data

Our source of data for adult outcomes is the American Community Survey (ACS). We merge the 2007-2013 surveys to form one set of repeated cross-sections. The ACS provides us with state and year of birth information for each respondent. To measure exposure to teacher performance pay programs, we calculate the fraction of students in each grade in each state and year who were enrolled in a school with performance pay. Assuming that a student begins 1st grade at age six and 12th grade at age seventeen, we take the summation of the fraction of students exposed over the 12 relevant years. Our measure could thus be thought of as the expected number of years of education a student would have had in schools where

[11]The low exposure we find in early in our sample is supported by the literature. Porwoll (1979) finds that only 4% of districts in a national survey provided any sort of merit pay to teachers in 1978, though it is not clear how frequently the merit pay was related to student test scores. In a qualitative study of 6 of the schools that Porwoll identified that still offered merit pay in 1983, Murnane and Cohen (1986) found that none related this pay to test scores, further supporting our low figures for this time period. A comprehensive survey by Cornett and Gaines (1994) suggests that test-based merit pay was virtually non-existent in this time-period and our data reflects that. Ballou (2001) reports that 10% of public school teachers in 1993-1994 were employed at a school which paid teachers merit pay, but again, it is not clear how frequently the merit pay is tied to student test scores. To our knowledge, our dataset is the first that is able to track merit pay programs that are explicitly linked to student test scores.

part of teacher pay was tied to student test scores.[12]

While we know the birthplace of each individual in our sample, we do not know whether they were educated in the state they were born. Similar to Loeb and Page (?), we use two samples. Our main sample will be students who were living in the same state as they were born one year before the survey. This sample will have a higher probability of receiving the exposure we calculate for them, but could produce negatively biased results if performance pay causes individuals to move to higher wage states. We will test the robustness of our results using a second sample which includes all individuals regardless of current residence. These results will be attenuated, as the uncertain location of education introduces measurement error into our exposure variable.[13]

We restrict attention to cohorts born after 1980, as earlier cohorts will have begun schooling before the CCD begins, and to individuals who were at least 23 years old at the time of the survey. We further restrict our sample to full-time, year round workers, though we include part-time and non-labor force participants in our robustness analysis. In Table 1, we display descriptive statistics for our ACS sample. Column (1) shows estimates for the sample that includes only those who still reside in their birth state. Performance pay exposure is small. The average individual in our sample was exposed to .74 years of schooling under a performance pay regime. In columns (2) and (3) we break this sample down between individuals who received zero and positive exposure, respectively. Conditional on being in a state-cohort which had any performance pay, the average individual receives 1.26 years of schooling at a school with performance pay. Cohorts with high performance pay exposure have a higher fraction of minorities, particularly Hispanics, reflecting the prevalence of these programs in the American Southwest. They appear to have slightly lower wages and education, and are in school districts which have slightly higher average per pupil spending.[14] Columns (4)

[12]Some students drop out of high school and never reach the 12th grade. We do not adjust the performance pay exposure measure for this as it is a possible outcome of the program.

[13]These results could also be biased upwards if the introduction of teacher performance pay causes families to move to states with better education systems that do not have such programs.

[14]This variable is also taken from the CCD and is measured only at the state-level. We interpolate missing nominal values and then convert them to real values using the CPI.

through (6) repeat this analysis for the sample which includes out-of-state movers, and we observe similar patterns.

IV Estimation Strategy

Our primary objective is to estimate the effect of an additional year of exposure to a teacher incentive pay program on adult wages. We do this by estimating the following regression equation

$$\ln\left(\text{wage}\right)_{ijtsy} = \beta_0 + \beta_1 \,\text{exposure}_{sy} + \beta_2 \,\text{spending}_{sy} + \gamma X_{ijtsy} + \theta_s + \delta_t + \gamma_y + u_{ijtsy} \qquad (5)$$

where the dependent variable is the log wage for worker i who is living in state j in year t and was born in state s in year y. The exposure variable is a continuous measure of the fraction of students born in state s in year y who were in a school with a teacher performance pay program summed up over grades 1 through 12. Therefore, the state and cohort-specific exposure variable has a minimum value of zero, a maximum value of 12, and can take on any value in between. In all specifications we include year fixed effects, δ_t to control for wage inflation and national employment shocks that may be correlated with performance pay programs. To control for age and cohort effects we include birth year fixed effects, γ_y.

In some specifications we include a state and cohort-specific measure of school spending per student averaged over the years in grades 1 through 12. We include birth state fixed effects, θ_s, in some specifications but do not include fixed effects for the current state of residence, j, because of selection concerns about the type of person who moves away from their birth state, s. Some specifications include a set of worker-specific characteristics, X_i, including gender, race, ethnicity, and age squared.[15] Note that we do not include education in our set of worker-specific controls, as educational attainment is a potential outcome of

[15] The worker's age is captured by the year and birth-year fixed effects. We include age squared to capture the rapid returns to labor market experience in the mid-20's.

performance pay. In some specifications we also include birth state-specific linear time trends.

As shown in Figure 2, southern states are more likely to implement a teacher performance pay program than states in other regions. This fact alone implies that the raw correlation between performance pay programs and wages is negative. Our birth state fixed effects and worker characteristics should significantly reduce this negative bias. The identification of the effect would then come from the timing of the adoption of these programs. Figure 2 also shows that there is substantial heterogeneity in when each state adopts a performance pay program. A concern for the identification is that states which have a budgetary surplus are probably more likely to implement a performance pay program. To account for this we include a measure of school spending, time fixed effects, as well a birth state-specific time trends. The measure of per pupil spending is a particularly helpful control variable because it allows the regression to compare workers with the same level of per pupil spending where some of that spending came in the form of teacher performance pay for those with exposure to the program.

Equation 5 allows us to answer the causal question of how teacher performance pay programs as implemented in the United States (with all their imperfections) affect the future wages of students taught by teachers who are given a financial reward (of various unobserved strength) if their students perform well on an exam. Holding per pupil spending constant, where this spending total includes the teacher performance pay bonuses, means that the alternative to a performance pay program is an equivalent increase in school funding with no teacher performance pay. If money spent on teacher performance pay programs is more effective at helping students develop skills that are valued in the job market than other spending on education, our results will show a positive effect of exposure on wages. Alternatively, our results will show a negative effect if money spent on teacher performance pay programs is less effective on average than other spending on education.

To look for other impacts of teacher performance pay programs, we estimate alternative versions of equation 5 with labor market indicators and educational attainment indicators as the dependent variable rather than log wage. To look for heterogeneous effects, we estimate equation 5 separately by ability and demographic groups. To investigate the timing of the effect, we estimate a version of equation 5 which allows the exposure effect in high school to differ from that in grade school and middle school.

Finally, the test score literature has found important heterogeneity in the effectiveness of performance pay programs. Poorly-designed programs have been found to have no effect even on the test used in determining teacher bonus payments. We model the heterogeneity in program quality with a state-specific random coefficient model

$$\ln(\text{wage})_{ijtsy} = \beta_0 + \boldsymbol{\beta}_s' \,\text{exposure}_{sy} + \beta_2 \,\text{spending}_{sy} + \gamma X_{ijtsy} + \theta_s + \delta_t + \gamma_y + u_{ijtsy} \tag{6}$$

where $\beta_s \sim N(\beta, \sigma^2)$ is drawn from a normal distribution with mean β and variance σ^2 for each state. This specification allows the effect of teacher performance pay programs to differ by state (though not across time or for different programs within the same state) while only requiring the estimation of the β and σ^2 parameters. We estimate the random coefficients model using a two-step procedure suggested by Swamy (1970) in which the β_s for each of the 27 states that ever had performance pay programs are estimated in separate state-specific regressions. The β and σ^2 parameters of the normal distribution are then estimated from $\hat{\boldsymbol{\beta}}_{\boldsymbol{s}}$ using the inverse of the squared standard errors as weights. This method has the advantage of being a transparent and simple way to estimate the variation in the wage effect of performance pay programs.

V Results

Results from the estimation of equation 5 on our sample of individuals who remained in their birth states are shown in Table 2. With no additional controls other than observation year fixed effects, the estimated effect implies that a one year increase in effective exposure to a performance pay program is associated with a 2.4 percent reduction in wages. However, this reflects both that poorer performing school systems were more likely to implement performance pay and that these programs were concentrated in the lower-wage South. When we include birth year and birth state fixed effects along with a basic set of worker characteristics in column (2), the effect goes to almost exactly zero.

The simple fixed effect regression neglects that states may be on different economic trajectories and these may influence the decision to implement performance pay. For instance, states may choose to invest in new teaching programs, including expensive performance pay, when they experience robust economic growth and thus increasing tax revenue. When we include state-specific linear time-trends our estimate as reported in column (3) becomes negative and statistically significant. In other words, birth cohorts with high-levels of performance pay deviate negatively from state time-trends in wages.

In column (4) we include controls for the mean log of per pupil school spending in the individual's birth cohort (assuming twelve years of education in the birth state), the mean student-teacher ratio, the likely exposure to charter schools, and the number of years of their education in which the Republican party controlled the governership and the state legislature. The estimated effect implies that a one year increase in effective exposure is associated with a 0.9 percent reduction in wages. This suggests that a student who would be exposed to a full 12 years of performance pay would see a reduction in wages of 10.8 percent. This reduction in wages is equivalent to losing a little more than a year of education in a non-performance pay environment. This result suggests that teacher performance pay programs in the United States had a negative effect on wages on average. Columns (5) and (6) report the results using the full sample of workers, including those who have moved from their birth state, and the results are similar.

There is no statistically significant effect on other labor market outcomes for both workers who reside in their birth state and the full sample as shown in Table 3. The estimate in column (1) suggests that workers who were exposed to a performance pay program face a higher risk of unemployment. However, columns (2) and (5) show that this correlation is not robust to the inclusion of the full set of controls. Similarly, there is no statistically significant effect on labor force participation or full-time employment.

Table 4 presents the estimated effect of performance pay programs on education outcomes. Columns (2) through (4) indicate that there is no statistically significant effect on educational attainment for the sample that reside in their birth state. There is some evidence of a reduction in educational attainment when the sample of workers who no longer reside in their birth state are added to the regression in columns (5) through (7). In particular, students appear to be less likely to complete a four year degree.

As the results for individuals who reside in their birth state and the full sample are similar, we will hereafter (in Tables 5, 6, and 7) focus only on those who reside in their birth state. We feel this provides a cleaner natural experiment where we are not concerned about potential bias if performance pay programs influence which type of student moves. Results using the full sample are reported in the appendix as Tables A-5, A-6, and A-7.

Our earlier model showed that teachers may respond to performance pay by redirecting effort from high ability individuals to low ability individuals. As the ACS does not provide a direct measure of ability, we instead look at outcomes separately by educational attainment. This would of course be problematic if performance pay programs led to increased education attainment. For example, high school graduates may have lower wages because the implementation of teacher incentives lead to more low ability students graduating from high school. However, Table 4 shows that this is not the case. If anything, these programs appear to have a negative effect on education, meaning our estimates would be biased upwards for each group.

Table 5 shows that there is substantial effect heterogeneity by student ability. There is a large positive, though not statistically significant, effect for students who drop out of high school as shown in column (1). The estimated effect is negative for the 95 percent of workers with at least a high school education as shown in columns (2) through (4). The largest negative effect is for the most educated group of college graduates and is statistically significant. This suggests that students at the bottom of the ability distribution benefit from teacher performance pay as teachers focus on helping them improve their test scores. The cost of this refocusing of teacher attention is lower future wages for the rest of the students on average, particularly those at the top of the ability distribution.

In Table 6 we break down our results by demographic groups. The negative adult wage effects are driven primarily by female students. There effects for white and Hispanic male students are close to zero, with a positive estimated effect for black male students. This is again consistent with our model of teacher attention, suggesting that teachers shift focus from female students towards black male students. Our estimates suggest that Hispanic female students are most disadvantaged by exposure to teacher performance pay.

In all our specifications to this point, we have implicitly assumed that effect of exposure is the same regardless of in which grade the child is exposed. There are many reasons to think this is not the case. The child development literature, for instance, has found that interventions earlier in life are more effective than those later in life (e.g. Cunha, Heckman, and Schennach, 2010). To test this assumption, we split the exposure measure into three periods: grade school (grades 1-5), middle school (grades 6-8), and high school (grades 9-12). Table 7 reports results by completed years of education. We see some evidence for college graduates that the largest (detrimental) effect comes from exposure in grade school. However, there is no clear pattern among the other education groups, and in no case can we reject equality of coefficients across the different exposures.

Estimation of the random coefficient model, where state-specific effects are drawn from a normal distribution, yields estimates of the two parameters of the normal distribution. In a specification that controls for year fixed effects, worker characteristics, school inputs, birth year and birth state fixed effects and birth state linear time trends, we estimate that $\beta_s \sim N(-.0229, .0171)$. The point estimate for β is more negative than we found in Table 2, though we find it less convincing given the normality assumption used in the estimation method. However, if we think of each performance pay program in the United States as being drawn from our estimated normal distribution, the results suggest that 9 percent of performance pay programs cause wage increases holding per pupil spending constant. Our view is that the best teacher performance pay programs likely had positive wage effects on average, with the largest benefits going to low-ability, black, and male students. However, the results suggest that teacher performance pay programs in the United States have not historically been of this "best" type and tend to have underperformed the wage gains that would have been achieved had schools simply been given the money that funded these programs.

VI Conclusion

In this paper we estimate the effect of teacher performance pay programs in the United States on adult outcomes, primarily wages but also other labor market outcomes and educational attainment. This is in contrast to most of the literature which has evaluated teacher performance pay through student test scores which can be manipulated and do not capture important skills that are difficult to test but are important in determining wages.

We construct a comprehensive data set of schools which have implemented teacher performance pay programs across the United States since 1986. This enables us to calculate the fraction of students in each grade in each state who attend a school in which the teachers have the ability to receive additional pay based on student performance where student performance is generally measured by standardized test scores. We believe our data is the most comprehensive such data set ever constructed.

We find that cohorts with a larger share of individuals enrolled in schools with performance pay programs earn lower wages as adults when controlling for per pupil spending. This result is robust to including individual and school control variables as well as birth state and birth year fixed effects and birth state linear time trends. Across the distribution of student ability, it is the lowest-ability students who experience higher adult wages as a result of exposure to a teacher performance pay program, while high-ability students are most harmed. We hypothesize that this is the result of a shift in teacher attention away from higher-ability students toward the lower-ability students. Similarly we find some evidence that teachers are shifting attention toward black and male students. Whether or not this is socially desirable depends on the chosen welfare function as there are far more workers where the wage effect is negative than positive.

References

Ballou, D., "Pay for Performance in Public and Private Schools," *Economics of Education Review*, 20 (2001), 51–61.

Bond, T. N., and K. Lang, "The Evolution of the Black-White Test Score Gap in Grades K-3: The Fragility of Results," *The Review of Economics and Statistics*, 95 (2013), 1468–1479.

Card, D., and A. B. Krueger, "School Resources and Student Outcomes," *The Annals of the American Academy of Political and Social Science*, 559 (1998), 39–53.

Chetty, R., J. N. Friedman, and J. E. Rockoff, "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood," *The American Economic Review*, 104 (2014), 2633–2679.

Cooper, S. T., and E. Cohn, "Estimation of a frontier production function for the South Carolina educational process," *Economics of Education Review*, 16 (1997), 313–327.

Cornett, L. M., and G. F. Gaines, "Reflecting on Ten Years of Incentive Programs: The 1993 SREB Career Ladder Clearning House Survey," Institute of Education Sciences ERIC No. ED378163 (1994).

Cornwell, C., D. B. Mustard, and J. V. Parys, "Noncognitive Skills and the Gender Disparities in Test Scores and Teacher AAssessment: Evidence from Primary School," *Journal of Human Resources*, 48 (2013), 236–264.

Cunha, F., J. J. Heckman, and S. M. Schennach, "Estimating the Technology of Cognivie and Noncognitive Skill Formation," *Econometrica*, 78 (2010), 883–931.

Dee, T. S., "Teachers, Race, and Student Achievement in a Randomized Experiment," *The Review of Economics and Statistics*, 86 (2004), 195–210.

——— , "A Teacher like Me: Does Race, Ethnicity, or Gender Matter?," *The American Economic Review*, 95 (2005), 158–165.

Dee, T. S., and B. J. Keys, “Does merit pay reward good teachers? Evidence from a randomized experiment,” *Journal of Poli*, 23 (2004), 471–488.

Dobbie, W., and R. G. Fryer, “Charter Schools and Labor Market Outcomes,” Working Paper (2016).

Eberts, R., K. Hollenbeck, and J. Stone, “Teacher Performance Incentives and Student Outcomes,” *Journal of Human Resources*, 37 (2002), 913–927.

Fryer, R. G., “Teacher Incentives and Student Achievement: Evidence from New York City Public Schools,” *Journal of Labor Economics*, 31 (2013), 373–407.

Goodman, S. F., and L. J. Turner, “The Design of Teacher Incentive Pay and Educational Outcomes: Evidence from the New York City Bonus Program,” *Journal of Labor Economics* (2013).

Hanushek, E. A., “The Failure of Input-Based Resource Policies,” *Economic Journal*, 113 (2003), F64–F98.

Heckman, J., R. Pinto, and P. Savelyev, “UnderOutcomes the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes,” *The American Economic Review*, 103 (2013), 1–35.

Imberman, S. A., and M. F. Lovenheim, “Incentive Strength and Teacher Productivity: Evidence From a Group-Based Teacher Incentive Pay System,” *The Review of Economics and Statistics*, 97 (2015), 364–386.

Jackson, C. K., R. C. Johnson, and C. Persico, “The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms,” *The Quarterly Journal of Economics*, 131 (2016), 157–218.

Jacob, B. A., “Where the boys aren’t: Non-cognitive skills, returns to school and the gender gap in higher education,” *Economics of Education Review*, 21 (2002), 589–598.

Jacob, B. A., and S. D. Levitt, “Rotten Apples: An Investigation of the Prevalance of Predictors of Teacher Cheating,” *Quarterly Journal of Economics*, 118 (2003), 843–877.

Koretz, D., and S. Barron, (1998) *The Validity of Gains in Scores on the Kentucky Instructional Results Information System (KIRIS).* RAND.

Ladd, H. F., "The Dallas School Accountability and Incentive Program: an Evaluation of its Impact on Student Outcomes," *Economics of Education Review*, 18 (1999), 1–16.

Lavy, V., "Teachers' Pay for Performance in the Long-Run: Effects on Students' Educational and Labor Market Outcomes in Adulthood," NBER Working Paper No. 20983 (2015).

Lincove, J. A., "Can Teacher Incentive Pay Improve Student Performance on Standardized Tests?," Working Paper (2012).

Murnane, R. J., and D. K. Cohen, "Merit Pay and the Evaluation Problem," *Harvard Educational Review* (1986).

Neal, D., (2011) *Handbook of the Economics of Education*chap. The Design of Performance Pay in Education, 495–550. Elsevier.

——— , "The Consequences of Using One Assessment System to Pursue Two Objectives," *Journal of Economic Education* (2013).

Podgursky, M., "Teams versus Bureaucracies: Personnel Policy, Wage-Setting, and Teacher Quality in Traditional Public, Charter, and Private Schools," in *Charter School Outcomes*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc. (2007).

Podgursky, M., and M. Springer, "Teacher Performance Pay: A Review," *Journal of Policy Analysis and Management*, 26 (2007), 909–950.

Porwoll, P. J., (1979) *Merit Pay for Teachers*. Arlington, Va. Educational Research Service.

Sojourner, A. J., E. Mykerezi, and K. L. West, "Teacher Pay Reform and Productivity: Panel Data Evidence from Adoptions of Q-Comp in Minnesota," *The Journal of Human Resources*, 49 (2014), 945–981.

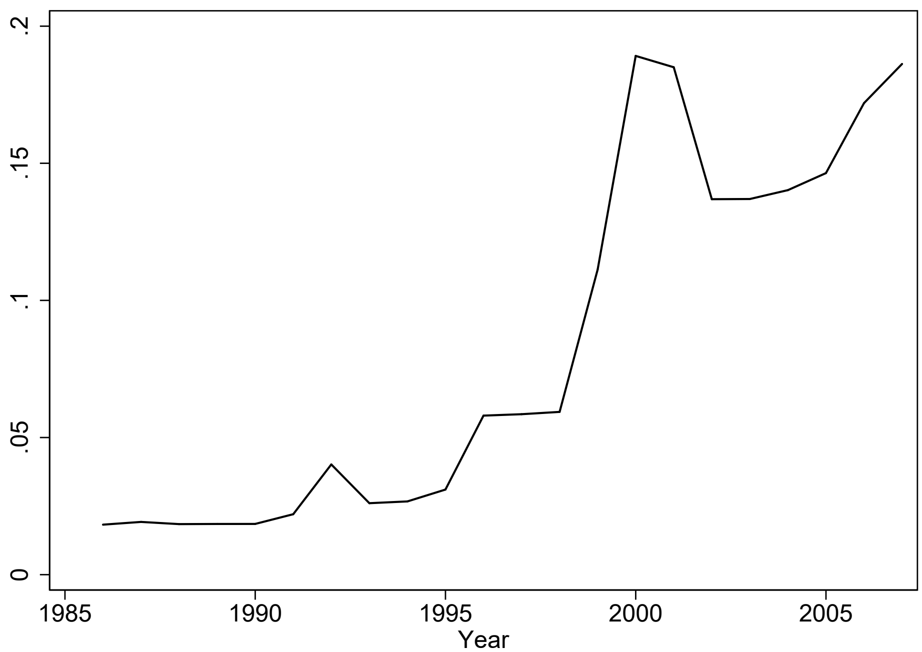
Swamy, P. A. V. B., "Efficient Inference in a Random Coefficient Regression Model," *Econometrica*, 38 (1970), 311–323.

Vigdor, J. L., (2009) *Performance Incentives: Their Growing Impact on American K-12 Education*chap. Teacher Salary Bonuses in North Carolina. Brookings.

Winters, M., J. P. Greene, G. Ritter, and R. Marsh, "The Effect of Performace-Pay in Little Rock, Arkansas on Student Achievement," National Center on Performace Incentives at Vanderbilt University Working Paper 2008-02 (2008).

Figures and Tables

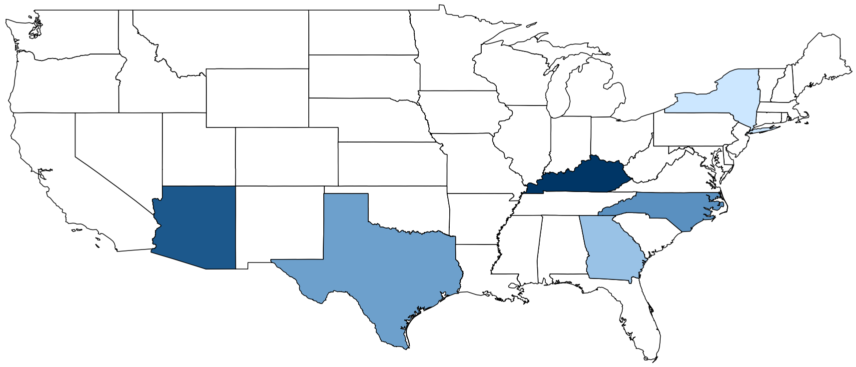
Figure 1: Fraction of Students in Schools with Performance Pay: 1986-2007



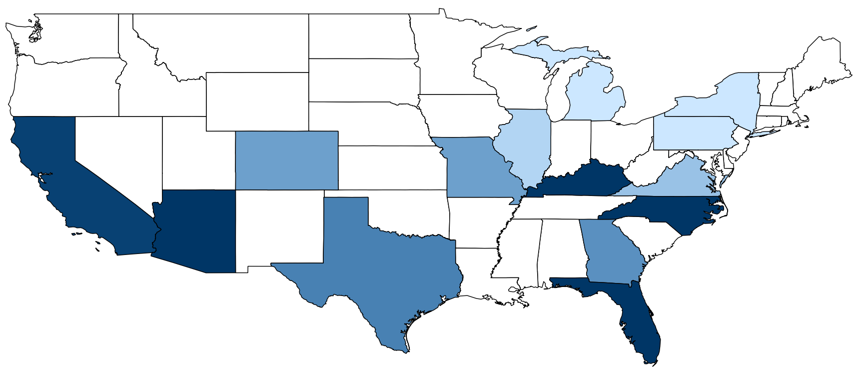
Source: authors construction of a panel dataset of teacher performance pay by grade in all non-charter schools in the U.S.

Figure 2: Fraction of Students in Schools with Performance Pay Program by State

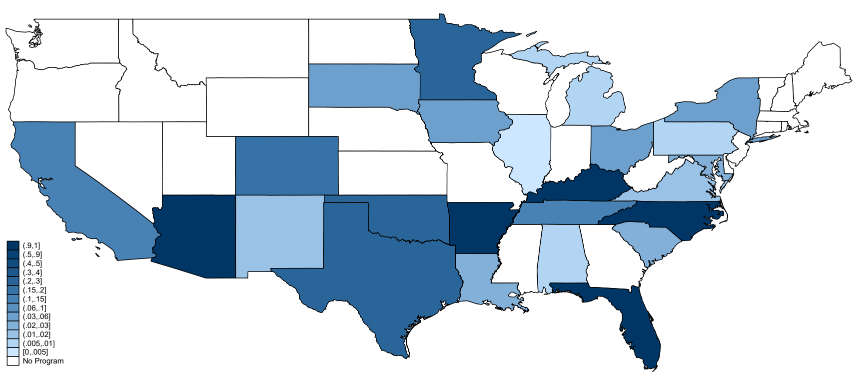
1993



2000



2007



Source: authors construction of a panel dataset of teacher performance pay programs by grade in all non-charter schools in the United States from 1986 to 2007. Note that Alaska and Hawaii are not shown. Alaska started a state-wide teacher performance pay program in 2006. Hawaii has no program over this time period.

Table 1: Descriptive Statistics

	Reside in Birth State			Full Sample		
	All	No Exposure	Positive Exposure	All	No Exposure	Positive Exposure
	(1)	(2)	(3)	(4)	(5)	(6)
Log Wage	2.70	2.75	2.66	2.72	2.78	2.68
	(0.54)	(0.53)	(0.54)	(0.55)	(0.55)	(0.55)
Exposure	0.74	0.00	1.26	0.70	0.00	1.25
	(1.91)	(0.00)	(2.35)	(1.87)	(0.00)	(2.35)
black	0.13	0.12	0.14	0.12	0.11	0.13
	(0.34)	(0.32)	(0.35)	(0.32)	(0.31)	(0.34)
hispanic	0.13	0.08	0.16	0.11	0.07	0.14
	(0.33)	(0.27)	(0.37)	(0.32)	(0.26)	(0.35)
Education	13.56	13.60	13.53	13.76	13.85	13.70
	(2.14)	(2.15)	(2.13)	(2.23)	(2.26)	(2.20)
Full-Time	0.85	0.86	0.84	0.85	0.86	0.84
	(0.36)	(0.35)	(0.37)	(0.36)	(0.35)	(0.37)
Unemployed	0.11	0.10	0.12	0.10	0.09	0.11
	(0.31)	(0.30)	(0.32)	(0.30)	(0.29)	(0.31)
Log Mean School Spending	8.86	8.83	8.88	8.86	8.84	8.88
	(0.22)	(0.23)	(0.20)	(0.22)	(0.24)	(0.21)
Student/Teacher Ratio	17.57	17.65	17.52	17.51	17.53	17.50
	(2.52)	(2.71)	(2.38)	(2.50)	(2.67)	(2.35)
Observations	1,148,233	769,679	378,554	1,667,474	1,150,553	516,921

The data for adult outcomes if from the 2007-2013 American Community Survey (ACS) and contains wage, education level, state of residence, and state of birth. The table reports summary statistics for the full sample as well as for the sub-sample of individuals who live in the state where they were born. For each sample, summary statistics are reported for those with no exposure to a teacher performance pay program and for those with a positive level of likely exposure to a teacher performance pay program. The standard deviation is reported in parenthesis below the relevant sample mean.

Table 2: Effect of Teacher Performance Pay on Log Wage

	Reside in Birth State				Full Sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure	-0.024***	-0.000	-0.010*	-0.009*	-0.024***	-0.005
	(0.006)	(0.002)	(0.006)	(0.005)	(0.006)	(0.003)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Worker Characteristics	No	Yes	Yes	Yes	No	Yes
School Inputs	No	No	No	Yes	No	Yes
Birth Year FE	No	Yes	Yes	Yes	No	Yes
Birth State FE	No	Yes	Yes	Yes	No	Yes
Birth State Time Trends	No	No	Yes	Yes	No	Yes
Observations	568,130	568,130	568,130	568,130	833,978	833,978

This table reports results from the estimation of equation 5 on the sample of workers who reside in their birth state. All specifications include year fixed effects. Column (2) includes birth-year fixed effects which, together with year fixed effects, uniquely identify the worker's age. Birth state fixed effects and controls for worker characteristics including gender, race, ethnicity, and age squared are also included. Column (3) adds birth state-specific linear time-trends. Column (4) also controls for the mean of the log of per pupil school spending over grades 1 through 12 in the individual's birth cohort, the student teacher ratio, and exposure to charter school programs. Columns (5) and (6) repeat the analysis using the full sample of workers. Standard errors are clustered by birth state and are reported in parentheses below the point estimates: * $p < .10$, ** $p < .05$, *** $p < .01$

Table 3: Effect of Teacher Performance Pay on Labor Market Outcomes

	Reside in Birth State				Full Sample		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Unemployed	Unemployed	Labor Force	Full-Time	Unemployed	Labor Force	Full-Time
Exposure	0.004***	-0.001	-0.004	0.002	0.000	-0.002	0.001
	(0.001)	(0.003)	(0.004)	(0.004)	(0.002)	(0.003)	(0.003)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Worker Characteristics	No	Yes	Yes	Yes	Yes	Yes	Yes
School Inputs	No	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Birthstate FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Birthstate Time Trends	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	931,736	931,736	1,148,233	931,269	1,358,322	1,667,474	1,362,073

The first four columns of this table report results from the estimation of equation 5 on the sample of workers who reside in their birth state by education level. The final three columns report results from the full sample. The dependent variable in columns (1), (2), and (5) is an indicator for unemployment. The dependent variable in columns (3) and (6) is an indicator for participation in the labor force. The dependent variable in columns (4) and (7) is an indicator for being a full-time worker. Standard errors are clustered by birth state and are reported in parentheses below the point estimates: * p< .10, ** p< .05, *** p< .01

Table 4: Effect of Teacher Performance Pay on Education Outcomes

	Reside in Birth State				Full Sample		
	(1) Years of Education	(2) Years of Education	(3) High School Degree	(4) Bachelor's Degree	(5) Years of Education	(6) High School Degree	(7) Bachelor's Degree
Exposure	-0.053***	-0.012	0.001	-0.003	-0.017**	0.000	-0.004**
	(0.015)	(0.011)	(0.002)	(0.002)	(0.007)	(0.002)	(0.002)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Worker Characteristics	No	Yes	Yes	Yes	Yes	Yes	Yes
School Inputs	No	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Birthstate FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Birthstate Time Trends	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,137,219	1,137,219	1,137,219	1,137,219	1,652,828	1,652,828	1,652,828

The first four columns of this table report results from the estimation of equation 5 on the sample of workers who reside in their birth state by education level. The final three columns report results from the full sample. The dependent variable in columns (1), (2), and (5) is the years of education completed. The dependent variable in columns (3) and (6) is an indicator for graduation from high school. The dependent variable in columns (4) and (7) is an indicator for graduation from college. Standard errors are clustered by birth state and are reported in parentheses below the point estimates: * $p < .10$, ** $p < .05$, *** $p < .01$

Table 5: Effect of Teacher Performance Pay on Log Wage by Education Group

	HS Dropouts (1)	HS Grads (2)	Some College (3)	College Grads (4)
Exposure	0.043	-0.012	-0.006	-0.015**
	(0.027)	(0.008)	(0.007)	(0.006)
Year FE	Yes	Yes	Yes	Yes
Worker Characteristics	Yes	Yes	Yes	Yes
School Inputs	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes
Birthstate FE	Yes	Yes	Yes	Yes
Birthstate Time Trends	Yes	Yes	Yes	Yes
Observations	25,015	136,915	208,372	195,537

The dependent variable is the log wage and the sample consists of workers who reside in their birth state. All specifications include year fixed effects, worker characteristics, school inputs, birth year and birth state fixed effects, and birth-state linear time trends. The results for workers who did not complete high school are given in column (1), results for workers with exactly a high school education are given in column (2), and results for workers with additional years of education beyond high school but less than a college degree are given in column (3), and results for workers with a college degree or more are given in column (4). Standard errors are clustered by birth state and are reported in parentheses below the point estimates: * p< .10, ** p< .05, *** p< .01

Table 6: Effect of Teacher Performance Pay on Log Wage by Demographic Group

	White Male	White Female	Black Male	Black Female	Hispanic Male	Hispanic Female
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure	0.001	-0.017	0.012	-0.002	-0.001	-0.035***
	(0.004)	(0.014)	(0.022)	(0.019)	(0.014)	(0.013)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Worker Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
School Inputs	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Birthstate FE	Yes	Yes	Yes	Yes	Yes	Yes
Birthstate Time Trends	Yes	Yes	Yes	Yes	Yes	Yes
Observations	221,984	194,964	23,667	32,188	35,500	33,061

This table reports the effects of exposure to a teacher performance pay program on log wage by demographic groups for workers who reside in their birth state. All specifications include year fixed effects, worker characteristics, school inputs, birth year and birth state fixed effects, and birth-state linear time trends. Standard errors are clustered by birth state and are reported in parentheses below the point estimates: * p< .10, ** p< .05, *** p< .01

Table 7: Heterogeneous Effect of Teacher Performance Pay on Log Wage by Grade and Education Group

	HS Dropouts (1)	HS Grads (2)	Some College (3)	College Grads (4)
Grade School Exposure	0.043	-0.012	0.001	-0.021**
	(0.027)	(0.011)	(0.006)	(0.009)
Middle School Exposure	0.034	-0.002	-0.018**	-0.018**
	(0.031)	(0.008)	(0.008)	(0.009)
High School Exposure	0.047*	-0.015**	-0.005	-0.012*
	(0.026)	(0.007)	(0.005)	(0.006)
Year FE	Yes	Yes	Yes	Yes
Worker Characteristics	Yes	Yes	Yes	Yes
School Inputs	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes
Birthstate FE	Yes	Yes	Yes	Yes
Birthstate Time Trends	Yes	Yes	Yes	Yes
Observations	25,015	136,915	208,372	195,537

This table reports the effects of exposure to a teacher performance pay program by demographic groups on log wage where the measure of exposure is split into three periods: grade school, middle school, and high school. All specifications include year fixed effects, worker characteristics, school inputs, birth year and birth state fixed effects, and birth-state linear time trends. The results for workers who did not complete high school are given in column (1), results for workers with exactly a high school education are given in column (2), and results for workers with additional years of education beyond high school but less than a college degree are given in column (3), and results for workers with a college degree or more are given in column (4). Standard errors are clustered by birth state and are reported in parentheses below the point estimates: * p< .10, ** p< .05, *** p< .01

Table A-5: Full Sample Results: Effect of Teacher Performance Pay on Log Wage by Education Group

	HS Dropouts	HS Grads	Some College	College Grads
	(1)	(2)	(3)	(4)
Exposure	0.036	-0.003	-0.001	-0.010*
	(0.023)	(0.006)	(0.005)	(0.006)
Year FE	Yes	Yes	Yes	Yes
Worker Characteristics	Yes	Yes	Yes	Yes
School Inputs	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes
Birthstate FE	Yes	Yes	Yes	Yes
Birthstate Time Trends	Yes	Yes	Yes	Yes
Observations	33,143	182,619	293,168	321,860

The dependent variable is the log wage. All specifications include year fixed effects, worker characteristics, school inputs, birth year and birth state fixed effects, and birth-state linear time trends. The results for workers who did not complete high school are given in column (1), results for workers with exactly a high school education are given in column (2), and results for workers with additional years of education beyond high school but less than a college degree are given in column (3), and results for workers with a college degree or more are given in column (4). Standard errors are clustered by birth state and are reported in parentheses below the point estimates: * p< .10, ** p< .05, *** p< .01

Table A-6: Full Sample Results: Effect of Teacher Performance Pay on Log Wage by Demographic Group

	White Male (1)	White Female (2)	Black Male (3)	Black Female (4)	Hispanic Male (5)	Hispanic Female (6)
Exposure	0.003	-0.011	0.006	0.004	-0.004	-0.020*
	(0.004)	(0.010)	(0.021)	(0.016)	(0.011)	(0.011)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Worker Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
School Inputs	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Birthstate FE	Yes	Yes	Yes	Yes	Yes	Yes
Birthstate Time Trends	Yes	Yes	Yes	Yes	Yes	Yes
Observations	332,162	291,667	34,323	43,587	47,718	43,224

This table reports the effects of exposure to a teacher performance pay program on log wage by demographic groups. All specifications include year fixed effects, worker characteristics, school inputs, birth year and birth state fixed effects, and birth-state linear time trends. Standard errors are clustered by birth state and are reported in parentheses below the point estimates: * p< .10, ** p< .05, *** p< .01

Table A-7: Full Sample Results: Heterogeneous Effect of Teacher Performance Pay on Log Wage by Grade and Education Group

	HS Dropouts	HS Grads	Some College	College Grads
	(1)	(2)	(3)	(4)
Grade School Exposure	0.033	-0.008	0.010*	-0.017*
	(0.023)	(0.008)	(0.006)	(0.009)
Middle School Exposure	0.037*	0.000	-0.009	-0.016**
	(0.022)	(0.009)	(0.008)	(0.007)
High School Exposure	0.036	-0.002	-0.003	-0.006
	(0.024)	(0.006)	(0.006)	(0.006)
Year FE	Yes	Yes	Yes	Yes
Worker Characteristics	Yes	Yes	Yes	Yes
School Inputs	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes
Birthstate FE	Yes	Yes	Yes	Yes
Birthstate Time Trends	Yes	Yes	Yes	Yes
Observations	33,143	182,619	293,168	321,860

This table reports the effects of exposure to a teacher performance pay program by demographic groups on log wage where the measure of exposure is split into three periods: grade school, middle school, and high school. All specifications include year fixed effects, worker characteristics, school inputs, birth year and birth state fixed effects, and birth-state linear time trends. The results for workers who did not complete high school are given in column (1), results for workers with exactly a high school education are given in column (2), and results for workers with additional years of education beyond high school but less than a college degree are given in column (3), and results for workers with a college degree or more are given in column (4). Standard errors are clustered by birth state and are reported in parentheses below the point estimates: * p< .10, ** p< .05, *** p< .01