Signals of concern about teacher working conditions have become a new norm in the public and policy spheres in the United States. These signals—especially those expressed in recent waves of teacher strikes and protests episodes—have increased the pressure to improve teacher pay and working conditions. Unfortunately, there is a dearth of evidence to help understand what the ultimate consequences of teacher salary increases are, not just for teachers but also for public educational systems. State legislatures that are considering approving salary increases for teachers, as well as federal policymakers who are discussing similar initiatives are in need of more research-based guidance to help them make meaningful proposals.

Surprisingly, there is very limited evidence on the relationship between teacher pay levels and student performance, especially from studies that either rely on national level data and/or use proper quasi-experimental methods (Britton & Propper, 2016; Figlio & Kenny, 2007; Loeb & Page, 2000). Similarly, little is known regarding how teacher wage raises may have an impact on any of the persistent achievement gaps and education inequities in the country (i.e., race, socioeconomic status). Both facts are in part driven by lack of nationally representative microdata that includes schools’ personnel practices in conjunction with, or linkable with, information of student performance at the district level. Therefore, the lack of reliable, representative, and updated evidence certainly inhibits knowing the role teacher salary plays as a factor that improves the quality of the education system and its productivity.

Our study examines the following research questions:
what is the relationship between school teacher pay and students’ test scores in US districts? How does the relationship between the two vary across different subgroups of students? To answer these questions, we use two sets of nationally representative data. We benefit from the existence of district level performance data for the U.S. school districts (from the Stanford Education Data Archive, between 2009 and 2015), which we combine with teacher salary information at the district level (from the National Center for Education Statistics (2007–2008) and National Center for Education Statistics (2011–2012) Schools and Staffing Survey and the 2015–2016 National Teacher and Principal Survey).

This offers us the possibility to exploit within-state and within-year variation in teacher salaries and performance outcomes. State fixed effects, however, absorb a substantial amount of the variation in both student achievement and teacher salaries. Because much of the variation in teacher salaries is across states rather than within states, we also employ a multilevel (hierarchical) mixed effects model, where each state is a higher-level entity and every school

---

Emma García1 and Eunice S. Han2

**Abstract**
This paper examines the relationship between teacher pay and students’ academic achievement, using nationally representative, district-level linked data between districts’ performance on standardized tests and average teacher base salary. By employing state fixed effects and multilevel mixed effects models, we find that both mathematics and English test scores are significantly higher in districts that offer higher base salaries to teachers, compared with those in districts with a lower teacher base salary. We also find that higher teacher base salaries reduce achievement gap between white and black students, as well as between white and Hispanic students, by raising test scores more for those minority students.

**Keywords**
teacher salary, NAEP, test scores, district performance, educational inequality

---

1Economic Policy Institute, Washington, DC, USA
2University of Utah, Salt Lake City, USA

**Corresponding Author:**
Eunice S. Han, Economics Department, University of Utah, 260 Central Campus Drive, Gardner Commons, Rm 4100, Salt Lake City, UT 84112, USA.
Email: eunice.han@economics.utah.edu

---

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
district is a lower-level unit, to complement the fixed effects models. We control for a broad range of characteristics of the school districts and their communities, that may correlate with student performance, which help us reduce omitted variable bias in examining the association between teacher pay and student performance (Britton & Propper, 2016; Loeb & Page, 2000).

To our knowledge, our analysis is the first to model the relationship between teacher pay and student test scores at the district level, which is the ideal level of aggregation (as discussed by Loeb & Page, 2000 and by Figlio & Kenny, 2007), for the country. Because the availability of performance data for different subjects and grade-levels, the analyses allow us to ascertain if the patterns are consistent across them. Importantly, given the availability of performance data for different population subgroups—race/ethnic groups and socioeconomic level groups—our study also speaks to how changes in teacher salaries can contribute to closing the persistent achievement gaps.1

This research improves upon previous literature by making several important contributions. First, we rely on nationally representative data that allow us to utilize the full breadth of variation of teacher salary and offer higher levels of external validity. Second, we use the base salary of teachers, instead of the fixed salary schedules of districts used in many studies, to better capture the variation of teacher compensation and produce more precise estimates. Third, we control for various teacher, district, and community characteristics, substantially reducing potential omitted variable bias. Finally, we conduct a comprehensive analysis of districts’ performance by disaggregating our results for various subgroups of students.

**Literature Review**

Two main approaches can be used to frame the empirical literature that has examined teacher wages and their relationship with teachers’ labor supply and student performance. One approach builds on ‘efficiency wages’ models used in labor economics to address issues of quality, effort, or productivity in specific labor markets. The second approach, borrowed from the economics of education, relies on traditional models of education production and how the allocation of resources to teacher-related inputs affects it. These two lines of thoughts may reveal important mechanisms through which higher salaries would help build a stronger teaching workforce (through influencing the pool of applicants, recruitment, and retention), which will ultimately improve student performance.

Efficiency wages are used in labor economics to argue that in the presence of different challenges—such as difficulty to identify quality, agency, effort, and labor outcomes—wages above market equilibrium can lead to increased labor productivity (Katz, 1986; Krueger and Summers, 1988; Stiglitz, 1986; Weiss, 1980, 2017, etc.). Higher salaries can help reduce costs associated with teacher turnover, and the increased productivity can pay for the higher salaries. This frame is also useful to examine the relationship between salaries and teacher quality and productivity—and hence, potentially student performance—because similar challenges can affect teacher labor markets. Measuring teachers’ productivity is affected by the difficulty to parcel out their contribution to an outcome that is jointly produced, by problems in the principal-agent relationships (aggravated by students, parents, and other agents in this relationship), or by monopsonistic, non-competitive features. Gius (2012) summarizes some of these challenges, building on existing research that argues about the difficulty to distinguish between individual levels of attainment “whenever outputs is jointly produced” (Murnane & Cohen, 1986), or that the principal-agent theory (specifically for merit pay purposes) is particularly difficult to implement in an education setting (Goldhaber et al., 2008). Other challenges appear because effort and quality are difficult to observe and measure, which has been used to call for the need to promote competition to increase productivity (Hanushek, 2015) because teacher labor markets may show some monopsonistic (Council of Economic Advisers [CEA], 2016).

Changes in salaries could lead to changes in quality and productivity through the following channels. First, the evidence demonstrates that higher salaries lead to higher quality students entering education, which would both strengthen the prospective teachers’ applicant pool and eventually expand the future teaching pipeline (Figlio, 1997; Hanushek et al., 2019; Leigh, 2012; Podolsky et al., 2019), where the quality of the potential pool of potential teachers is measured by teachers’ ability, test scores, their college selectivity, or other credentials. Along these lines, Figlio (1997) finds that higher wages increase the share of teachers who graduated from a selective college and the share of teachers with subject-matter qualification. For example, he finds that, in metropolitan areas, an increase of 1% in average teacher salaries would increase the share of teachers who have graduated from a selective college by 1.6%. The results hold both in the national level sample and when testing the relationship in six specific metropolitan areas. Leigh (2012) shows that an increase of 1% in the salary of starting teachers raises the average aptitude of students who enter teacher education courses by 0.6 percentile ranks. This research, which uses a model of current salaries and academic aptitude of potential future teachers and test scores for all students admitted to a university in Australia between 1989 and 2003, finds a stronger effect for those at the median of the distribution.

A related issue is that higher salaries correlate with increased interest in becoming teachers, which can potentially strengthen the quality of the pool of potential teachers. Early work by Manski (1987) shows that a 10% increase in teachers’ weekly salary increases the proportion of college graduates willing to work as teachers by 26%. This, in part, is in agreement with the lack of attractiveness of teaching...
as a career due to low pay among recent cohorts of students taking the ACT, who chose low salary as the most cited reason when asked why they were not interested in teaching (Croft et al., 2018).²

Relative pay for teachers can also influence who enters the teaching profession. In Florida, the effectiveness of teachers who entered the profession during recessions was higher than for those who did so during non-recessionary periods, with the former being about 0.10 standard deviations (0.04) more effective in raising math (reading) test scores (Nagler et al., 2017), whereas an economic boom in Texas (that tripled the local tax base and boosted revenues via shale oil and gas drilling) reduced test scores and student attendance (Marchand & Weber, 2019).

The positive association between teachers’ salary and their skills is also observed using international evidence. For example, Hanushek et al. (2019) rely on international data for over 30 OECD countries to show that countries that pay teachers more “tend to draw their teachers from higher parts of the college skill distribution” (p. 63).

The second channel is that higher salaries lead to higher quality teachers. In general, the labor economics literature finds a positive relationship between wages and skills (or positive returns to skills other than educational attainment as in Murnane et al., 1995; etc.). In education, higher salaries have been shown to increase average credentialing and experience of teachers (Hendricks, 2014; Ronfeldt et al., 2013; Sorensen & Ladd, 2018). The empirical evidence produced more recently counters a historical view claiming a lack of a systematic link between quality and salary and working conditions (Hanushek & Rivkin, 2007).

Some studies show that changes in teachers’ effectiveness or quality resulting from net turnover associated with changes in the relative pay for teachers (i.e., shocks to the labor market affect who stays in the profession) lead to changes in the quality of the teaching workforce. For example, Britton and Propper (2016) find that a 10% increase in the local labor market wage in England results in a 1.4 point decrease of teaching quality (a very large change relative to a mean of 2.5 points, where quality is measured by a metric produced by inspections by the national school regulator), an increase in the share of novice teachers, and a decrease of teachers with more than 10 years of experience. Hendricks (2016) finds that increasing salaries for teachers with three or more years of experience raises high-ability teachers’ retention, whereas higher salaries for teachers with 0 to 2 years of experience increases the retention rate of low-ability teachers, where ability is measured by certification scores, in Texas public schools during the 1995 to 1996 through 2013 to 2014 school years.

The third channel is that higher salaries reduce teacher turnover, or increase retention, which is helpful for students and schools. The relationship between wages and turnover is clear in multiple pieces of research. Wages are important for retaining and attracting teachers (Gray & Taie, 2015; Grissom et al., 2015; Katz, 2018; Loeb et al., 2005; Manski, 1987; Murnane & Olsen, 1989; Podolsky et al., 2019; Stockard & Lehman, 2004). In addition, salaries are particularly important for retention of teachers in their early careers, and in high-poverty or high-needs schools, where issues of turnover and quality are more striking (García & Weiss, 2019; Hanushek et al., 1999; Loeb et al., 2005; Sorensen & Ladd, 2018).

Hendricks (2014) finds that paying teachers leads to higher retention a year later, using data from Texas between 1996 and 2012. His study controls for a set of time-varying or fixed labor market conditions and district characteristics that could be correlated with teacher pay. He finds that a 1% increase in teacher pay reduces teacher turnover by 0.16 percentage points, with a larger effect for less experienced teachers. He simulates that, through this effect, paying teachers more improves student achievement (through higher retention, and through increasing the average experience of teachers in the district). With data from Texas as well, Hanushek et al. (1999) find that increasing teacher salaries within a district by 10% reduces a teacher’s probability to leave the district by 2% for probationary teachers, and by 1% for teachers with 3 to 5 years of experience.³

Even having more equal salaries between teachers and non-teachers reduces teacher turnover. On relative salaries or opportunity costs, Murnane and Olsen (1989) look into how the salary of the best job alternative outside of teaching affects teachers’ retention and find that $1,000 (in 1967 $) in the opportunity cost salary correlates with a decrease of 4 years in the median length of stay in teaching (using data from Michigan’s State Department of Education, which followed teachers who started their teaching careers in the early 70s until the 1984–1985 school year). A share of studies covered below incorporates the idea of the teacher salary gap and, thus, controls for this channel, into their empirical strategies.⁴

Research is also conclusive on how turnover and attrition affect student performance. A lack of sufficient, qualified teachers threatens students’ learning ability (Darling-Hammond, 2000; Ladd & Sorensen, 2016). Instability in the teaching workforce in a school due to high turnover or high attrition negatively affects student achievement, and it diminishes teacher effectiveness and quality (Darling-Hammond, 2000; Jackson & Bruegmann, 2009; Kraft & Papay, 2014; Ladd & Sorensen, 2016; Ronfeldt et al., 2013; Sorensen & Ladd, 2018). Turnover especially depresses student achievement in the highest-poverty schools, with “turnover-induced loss of general and grade-specific experience” as the main driver of declining student achievement (see Sorensen & Ladd, 2018, citing Hanushek et al., 2016). As mentioned earlier, net turnover increases the share of inexperienced teachers who are not fully certified or credentialed to teach the subject to which they are assigned. The turnover begets further turnover, which substantially weakens the overall quality and ability of the teacher pool in a given
school (Sorensen & Ladd, 2018). For novice teachers, Gray and Taie (2015) note that there is a 9 to 10 percentage-point gap in the rates of attrition between teachers who have a first-year-salary of $40,000 or more as compared with those earning less.5

The second approach relies on the traditional production of education models and resource allocation. From the perspective of the productivity of education spending, it is known that more resources increase student performance (Jackson, 2018; LaFortune et al., 2018), and that teachers are the most important performance factor within the school (Hanushek et al., 1998; Hartel, 2013). For these reasons, it would be expected that an increase of spending on teacher factors, and/or a reallocation of resources toward these factors would yield increases in student performance. This approach always keeps the focus on some metric of productivity or of student performance (for example, test scores, student learning, graduation rates, or some non-contemporaneous outcome linked to performance, such as labor income as adults, etc.) and uses some measurement of teacher compensation to examine their association.

The empirical evidence produced so far on this front was deemed as mixed by some (Hanushek, 1997, 2003, 2015), but also as almost always positive by others (Glewwee et al., 2014), and it has been accepted as limited, incomplete, often challenged by data limitations. Shedding some light on this puzzling summary, the departing point, still recently cited, was that “there is very weak support for the notion that simply providing higher teacher salaries or greater overall spending will lead to improved student performance” (Hanushek, 2015, p. 152, based on revised Hanushek, 1997, 2003).6 This statement is made out of meta-analysis results and the tabulation of estimates by their sign and statistical significance. The original study (Hanushek, 2003) reveals that 20% out of a total of 118 estimates showed a positive coefficient, 7% negative, and 73% were statistically insignificant.7

More recent evidence based on alternative methods that use variation in teacher pay from state-level or country-level data, and sometimes relies on natural experiments, shows that higher wages lead to positive student outcomes. (For case studies, usually covering one or a small number of school districts, see Lin, 2010.)

For example, Loeb and Page (2000) find that a 10% increase in wages reduces the dropout rate (the percentage of 16–19 year-olds who are not attending high-school and do not possess a high-school diploma) by 3% to 6% about 10 years later (the argument being that “it takes time for wage changes to lead to higher average teacher quality,” see p. 397). Using college attendance as an alternative outcome, the paper finds an increase of 1.6% in college enrollment for a 10% wage increase. This study uses state-level data and controls for both teacher and non-teacher salaries, and tests for the consistency of the findings using district level data.8 Card and Krueger (1992) find that, for white males born between 1920 and 1949, a 10% increase in teacher salaries led to a 0.1 percentage-point increase in the rate of return to schooling.9

Britton and Propper (2016) find that a 10% increase in the teacher pay penalty results in an average loss of about 2% in average school performance (high stakes tests taken at the end of compulsory schooling in England). Their identification strategy uses a natural experiment created by a characteristic of how teacher salaries are set in England (they are centrally regulated), which generates an exogenous gap between wages of teachers and of non-teachers.

Another source of growing descriptive or correlational evidence lies on cross-country analyses. Most of these analyses also show a positive association between teacher wages (at the country level) and student performance (Boarini & Lüdemann, 2009; Dolton & Marcenaro-Gutiérrez, 2011; Hanushek et al., 2019). Though their methods are inadequate to issue causal claims, analyses exploiting these different sources of variation have contributed to raising awareness of differentials in teacher base salaries and compensation across countries that could in part explain the systems’ aggregate performance.

Data

This study utilizes district-level linked data from the Schools and Staffing Survey (SASS) 2007 to 2008, the Schools and Staffing Survey (SASS) 2011 to 2012, the National Teacher and Principal Survey (NTPS) 2015 to 2016, and the Stanford Education Data Archive (SEDA) 2009, 2012, and 2015. The SASS and NTPS, administered by the National Center for Education Statistics (NCES), are nationally representative data that contain about a third of public school districts in the United States. Replacing the SASS beginning the 2015 to 2016 academic year, the NTPS is redesigned with a focus on “flexibility, timeliness, and integration with other Department of Education data” (NCES, 2015–2016). Both the NTPS and SASS include detailed questionnaires at the teacher, principal, and school level, while the SASS also contains school district information. We restrict our analyses to teachers in public schools.

The SEDA, administered by the Center for Education Policy Analysis at Stanford University, incorporates our main outcome variables for performance—district-level averages of students’ test scores, in mathematics and reading, for grades third to eighth. The SEDA also includes information on the characteristics of students and their families and schools at the district level that are based on the Department of Education’s Common Core Data (CCD) and the American Community Survey via the NCES School Districts Demographic System (SDDS).

We merge the SASS, NTPS, and SEDA to construct a data set containing detailed information on school districts, teachers, students, and their communities, at three points in time, based on a unique school district ID number. Our dataset thus...
consists of three waves of information on the included districts: district performance and characteristics from the 2008 to 2009 SEDA linked to teacher salary information from the 2007 to 2008 SASS, the 2011 to 2012 SEDA linked to the 2011 to 2012 SASS, and the 2014 to 2015 SEDA linked to the 2015 to 2016 NTPS (see Supplemental Appendix I for descriptive statistics). The total number of districts in our sample is approximately 10,000.

The SEDA contains student performance data from English and mathematics tests for the third through eighth grades. Using ordered probit models, the means and standard deviations of test scores are estimated based on reports of the counts of the number of students scoring in each proficiency category on these tests. Then, these means and standard deviations are converted to a common scale calibrated to the scores taken from the National Assessment of Educational Progress (NAEP) tests that are administered to fourth and eighth grade students in odd-numbered years (for a detailed discussion of these methods, see Fahle et al., 2018). Average means and standard deviations are reported by district, grade, year, and also separately by white, Black, Hispanic, and Asian students, as well as for students in different socio-economic status groups.

Our main variable of interest is teacher base salary, which comes from the SASS/NTPS. For each district, we compute the average of the base salary of individual teachers, weighted by each teacher’s final sample weight. We also compute the district-level averages for teacher’s characteristics, such as gender ratio, experience, certification status, union membership rate, and charter school enrollment, and use them as control variables.

We control for numerous characteristics of the districts and their neighborhoods and contrast districts within the same state, so that our results are based on the comparison between similar districts in various dimensions. We control for basic demographic characteristics of school districts including total grade school enrollment; the share of students who are Hispanic, Black, Asian, White, or Native American in each grade; the share of students that are English Language Learners in the district; the share of special education students in the district; the total number of teachers; the total number of instructional aides; the share of all students on free or reduced-price lunch programs; the share of public school students in charter schools; and the share of districts in an urban, suburban, town, or rural location. Additionally, we control for characteristics of the community because they are likely to be associated with districts’ socio-economic status (SES), which we measure with the share of children in poverty, median household income, the share of adults with a bachelor’s degree and above, the share of households with children and a female head, the share of residents living in the same house as in the prior year, the share of unemployed, and the Gini coefficient.

Figure 1 describes a relationship between teacher base salary and average math test scores for each district in 2015 to 2016 with a scatter plot and the line of best fit. The math test scores increase as teacher pay rises for both the fourth and eighth grades. Judging based on the steepness of the slopes of the fitted lines, the strength of the association between pay and district performance is greater for the eighth grade than for the fourth grade.

Figure 2 also presents the positive relationship between teacher pay and average English test scores of districts for the fourth and eighth grades. Both fourth and eighth grade students’ English performance improves as teachers receive a higher base salary. Unlike the math test scores, the slopes of the two lines are similar.

**Methods**

Although Figures 1 and 2 may suggest that teacher pay could improve overall district performance, they may depict a spurious correlation that occurs simply because students from affluent districts perform better than students from disadvantaged districts. The relationship between teacher pay and district performance may also be affected by other characteristics of the districts that may be correlated with both teacher pay and district performance.

To address this issue, we control for an extensive set of teacher, district, and community characteristics by incorporating the covariates, one category at a time. First, we add school district characteristics to control for factors that are directly associated with district performance. Second, we include community attributes as additional control variables to tease out their indirect effects on district characteristics and educational outcomes. Lastly, we control for average teacher characteristics within a district to further reduce potential omitted variable bias. See Supplemental Appendix I for summary statistics for all these variables by category and by test subject.

Even though we control for various characteristics of districts to minimize potential omitted variable bias, each state’s unique features may also influence teacher pay and student performance. For instance, each state has a different cultural and legal environment, as well as a unique accountability system for its schools, reflecting a general preference toward public education that is relatively constant over the years. Moreover, all districts are exposed to common shocks in any given year, and this time effect can vary by year, potentially biasing the results. Our first attempt to control for such confounding factors is to exploit within-state and within-year variation by estimating the following equation:

$$ Y_{kst} = \beta_0 + \beta_1 Salary_{kst} + \beta_2 X_{kst} + \delta_s + \lambda_t + \epsilon_{kst}, $$

where $k$, $s$, and $t$ indicate districts, states, and years, respectively. $Y_{kst}$ represents the test score, either mathematics or English. $Salary_{kst}$ measures the average teacher base salary in district $k$ in state $s$ in year $t$, and it is expressed in logarithms.
$X_{kr}$ is a vector of district and community characteristics. $\delta_s$ is the state dummy and $\lambda_t$ the year dummy. $\epsilon$ is the error term, reflecting variation not accounted for in the model.

Fixed effects models, however, absorb a substantial amount of the variation in both student achievement and teacher salaries in which much of the variation in teacher
salaries exists across states rather than within states. In addition, there may exist unobservable factors of districts that can pose an endogeneity problem. For instance, districts within the same state may share common characteristics and experiences, such as legal and cultural environment, that are unobservable. When this commonality is large, districts in the same state may not behave independently, as they have similar state-level residuals, and the standard OLS estimates will suffer from omitted variable bias.

To tackle these issues, we also employ the multilevel (hierarchical) linear model that separates the total variance into within-group and between-group components. We estimate the following multilevel mixed-effects linear model:

$$Y_{kt} = (β_0 + u_s) + β_1 Salary_{kt} + β_2 X_{kt} + λ_t + ε_{kt}, \quad (2)$$

where $k$, $s$, and $t$ indicate districts, states, and years, respectively. $X$ is the vector of control variables at the district level. Model (2) assumes that the relationship between teacher pay and district performance is the same ($β_s$) for all states, and it estimates a single coefficient for each independent variable (fixed effects). However, the model allows a state-specific intercept ($u_s$) for each state (random effects). Because the model has both fixed effects and random effects components, it is called “mixed-effect.” Supplemental Appendix II explains details on multilevel mixed-effects linear models in which the state is considered a higher level and the district a lower level.

Researchers find that there exist inequities in the access to teacher credentials and the impacts of resources differ by various racial/ethnic and SES groups (Adamson & Darling-Hammond, 2012; Clotfelter et al., 2006; Goldhaber et al., 2014; Isenberg et al., 2013; LaFortune et al., 2018; Sass et al., 2012). Building on these studies, we run separate analyses by grade, for different race and ethnicity groups of students, and for different poverty level groups. Each subgroup shares the same cultural values and philosophy, which tend to be stable but unobservable. Thus, this subgroup analysis can provide more reliable estimates, because it also controls for those unobservable factors.

**Results**

Table 1 presents the estimated results from state and year fixed effects for all students for mathematics test scores, pooling all grade and race-ethnicity groups together. All model specifications show significantly positive association between teacher base salary and districts’ math performance. In model (1), the correlation coefficient between the log of base salary and math test scores is about 10.5, indicating that a 10% increase in teacher base salary is associated with a 1.05 higher average math test score. When normalized, this is equivalent to about one-tenth of a standard deviation in district average math test scores. We control for district characteristics in model (2). The coefficient for base salary substantially falls to about 4, but it remains significant at the 1% significance level. After controlling for community characteristics in the regression in model (3), and adding average teacher attributes as additional control variables in model (4), the coefficients for teacher salary are cut in half, but the significance is still intact. A 10% increase in teacher salary is associated with about 0.2 points (0.01 of a standard deviation) higher average math score.

Table 2 shows similar results for English test scores. We find a statistically positive relationship between teacher pay and English test scores. The correlation coefficient in model (1) is 12, which is greater than that for math scores. However, model (4) with full sets of covariates shows almost identical results with math: a 10% increase in teacher salary is linked with a higher average English test score of about 0.2 points.

Tables 1 and 2 display that, compared to white students, Hispanic, Black, and Native American students score significantly worse on both math and ELA tests, whereas Asian students perform better. Districts with a higher proportion of English Language Learners (ELL), special education students, and students under free and reduced-price lunch programs tend to score poorly on standardized tests.

Because there may be a trade-off between hiring teachers and hiring instructional aides (paying teachers more while hiring fewer aides or hiring fewer teachers but more aides, etc.), we control for both the number of teachers and aides. The results show that hiring more instructional aides is associated with lower test scores, and this may be partly because some districts are required to hire fewer teachers or pay lower salaries to teachers to maintain their budgets.

Districts with a larger fraction of adults with a higher level of education tend to score higher on these tests, whereas districts with a greater fraction of female-headed households and higher unemployment perform worse.

For math test scores, compared to districts with more novice teachers (with 1–2 years of experience), districts with a greater fraction of early-career teachers (with 3–5 years of experience) perform better. For English, districts with a larger proportion of female teachers and experienced teachers (5 years and above) score significantly higher.

The test scores used in this study are estimated values (with known error variance) of the parameters from different sites, where the true values of those parameters are assumed to vary among those sites. Thus, as a sensitivity test, we also use a meta-analytic regression model to account for known precision in test score estimates (see Reardon et al., 2019 for detailed explanation on this approach). Overall, the alternative results are similar to the previous results.16

The estimates reported in Tables 1 and 2 solely rely on within-state variation, which is smaller than between-state variation. To capture the full breadth of the variation of teacher salary while still controlling for unobservable commonalities that are shared by districts within each state, we consider multilevel models.
Table 1. Estimated Relationship Between Teacher Pay and Districts’ Math Performance.

| Variables                                  | (1)            | (2)            | (3)            | (4)            |
|--------------------------------------------|----------------|----------------|----------------|----------------|
| Log (Base salary)                          | 10.525*** (0.836) | 3.910*** (0.461) | 1.737*** (0.419) | 1.827*** (0.478) |
| District characteristics                   |                |                |                |                |
| % Hispanic students                        | −2.883*** (1.020) | −4.593*** (0.923) | −4.559*** (0.922) |                |
| % Black students                           | −14.00*** (0.609) | −16.94*** (0.710) | −16.93*** (0.710) |                |
| % Asian students                           | 38.16*** (2.227) | 18.67*** (1.887) | 18.68*** (1.885) |                |
| % Native Americans                         | −12.90*** (1.475) | −14.54*** (1.359) | −14.50*** (1.359) |                |
| Total enrollment, grades 3–8 (in 1,000)    | −0.00342 (0.0288) | 0.0176 (0.0257) | 0.0174 (0.0257) |                |
| % ELL students                             | −3.759*** (1.795) | −4.286*** (1.735) | −4.332*** (1.735) |                |
| % special education students               | −7.150*** (2.211) | −6.108*** (1.672) | −6.203*** (1.676) |                |
| Total teachers (in 1,000)                  | 0.179 (0.184)    | −0.0334 (0.174) | −0.0327 (0.174) |                |
| Total instructional aides (in 1,000)       | −0.423 (0.327)   | −1.086** (0.442) | −1.087** (0.445) |                |
| % reduced lunch                            | −33.28*** (2.256) | −12.14*** (2.237) | −12.24*** (2.228) |                |
| % free lunch                               | −30.16*** (0.687) | −15.81*** (0.864) | −15.81*** (0.862) |                |
| Community characteristics                  |                |                |                |                |
| Log(Median household income)               | 1.627** (0.646)  | 1.593** (0.648) |                |                |
| % Adults with bachelor’s degree and above  | 24.52*** (1.001) | 24.50*** (1.000) |                |                |
| % 5–17 year olds in poverty                | 5.105*** (1.814) | 5.121*** (1.814) |                |                |
| % Household with children and female head  | −4.640*** (1.429) | −4.660*** (1.429) |                |                |
| % Unemployed                               | −20.17*** (4.991) | −20.48*** (5.002) |                |                |
| % Living in same house as last year        | 2.491* (1.343)  | 2.563* (1.344)  |                |                |
| Gini coefficient                           | −2.758 (2.470)  | −2.780 (2.472)  |                |                |
| City/urban locale                          | 0.133 (0.282)   | 0.105 (0.283)   |                |                |
| Suburban locale                            | 0.817*** (0.218) | 0.791*** (0.218) |                |                |
| Town locale                                | 0.679*** (0.190) | 0.671*** (0.191) |                |                |
| Teacher characteristics                    |                |                |                |                |
| % Female teachers                          | 0.384 (0.266)   |                |                |                |
| % of teachers with regular state certificate | 0.0212 (0.466)  |                |                |                |
| % of teachers with alternative certification | −0.103 (0.381) |                |                |                |
| % of teachers with 3–5 years of experience  | 1.106* (0.584)  |                |                |                |
| % of teachers with 6–20 years of experience | 0.513 (0.547)  |                |                |                |
| % of teachers with 21+ experience          | 0.456 (0.592)   |                |                |                |
| Fourth grade                               | 10.15*** (0.0715) | 10.16*** (0.0701) | 10.16*** (0.0701) |                |
| Fifth grade                                | 20.29*** (0.0879) | 20.31*** (0.0868) | 20.31*** (0.0868) |                |
| Sixth grade                                | 31.06*** (0.102) | 31.07*** (0.101) | 31.07*** (0.101) |                |
| Seventh grade                              | 41.43*** (0.117) | 41.44*** (0.117) | 41.44*** (0.117) |                |
| Eighth grade                               | 52.09*** (0.135) | 52.16*** (0.134) | 52.16*** (0.134) |                |
| Constant                                   | 154.7*** (7.737) | 201.3*** (5.104) | 196.0*** (8.565) | 194.5*** (8.711) |
| Year dummies                               | X              | X              | X              | X              |
| State dummies                              | X              | X              | X              | X              |
| Observations                               | 57,010         | 56,650         | 55,710         | 55,710         |
| \( R^2 \)                                  | .009           | .860           | .874           | .874           |

Source: Authors’ calculation based on 2008 to 2009, 2011 to 2012, and 2014 to 2015 Stanford Education Data Archive (SEDA, v. 2.1) combined with 2007 to 2008 and 2011 to 2012 School and Staffing Survey (SASS) and 2015 to 2016 National Teacher and Principal Survey (NTPS).

Note: Errors are clustered within states (presented in parentheses).

*\( p < .1 \), **\( p < .05 \), ***\( p < .01 \).

N is rounded to nearest 10.

First, to examine if the commonality among districts in the same state is large or small, we estimate intraclass correlation (\( \rho \)), which is a summary of the proportion of the outcome variability that is attributable to differences across states. Larger values of \( \rho \) (close to 1) imply that districts in the same state behave almost identically. On the other hand, smaller values of \( \rho \) (close to 0) signal that the districts in the same state are almost independent from each other, and simple OLS regression could suffice for the analysis.

In our models, we estimate the between-district variance component \( (\sigma_u^2) \) to be 15.36 and within-district variance component \( (\sigma_\varepsilon^2) \) to be 61.46, yielding the intra-class correlation \( (\rho) \) of .201 for math test scores. The \( \rho \) for English test scores is .189. The sizable values of \( \rho \) imply that districts
Table 2. Estimated Relationship Between Teacher Pay and Districts’ English Performance.

| Variables                          | (1)          | (2)          | (3)          | (4)          |
|------------------------------------|--------------|--------------|--------------|--------------|
| Log (Base salary)                  | 11.64***     | 4.455***     | 1.979***     | 1.808***     |
|                                    | (0.901)      | (0.469)      | (0.400)      | (0.453)      |
| District characteristics            |              |              |              |              |
| % Hispanic students                 | −4.921***    | −6.699***    | −18.03***    | −18.03***    |
|                                    | (1.101)      | (0.977)      | (0.681)      | (0.682)      |
| % Black students                   | −14.33***    | −18.13***    | −19.12***    | −19.12***    |
|                                    | (0.607)      | (1.466)      | (1.463)      |              |
| % Asian students                   | 35.41***     | 13.00***     | 13.10***     | 13.10***     |
|                                    | (2.249)      | (1.748)      | (1.740)      |              |
| % Native Americans                 | −17.30***    | −19.26***    | −19.26***    | −19.12***    |
|                                    | (1.668)      | (1.466)      | (1.463)      |              |
| Total enrollment, grades 3–8 (in 1,000) | −0.00189     | 0.0298       | 0.0297       | 0.0225       |
|                                    | (0.0267)     | (0.0224)     | (0.0225)     |              |
| % ELL students                     | −12.19***    | −14.99***    | −19.81***    | −19.81***    |
|                                    | (2.132)      | (2.006)      | (2.004)      |              |
| % Special education students       | −8.037***    | −3.456*      | −3.575***    | −3.575***    |
|                                    | (2.412)      | (1.803)      | (1.804)      |              |
| Total teachers (in 1,000)          | 0.141 (0.174)| −0.166 (0.150)| −0.166 (0.151)|            |
| Total instructional aides (in 1,000)| −0.385 (0.303)| −1.227*** (0.356)| −1.229*** (0.361)| |
| % Reduced lunch                    | −38.55***    | −14.36***    | −14.46***    | −14.46***    |
|                                    | (2.121)      | (2.114)      | (2.099)      |              |
| % Free lunch                       | −33.53***    | −16.35***    | −16.32***    | −16.32***    |
|                                    | (0.720)      | (0.882)      | (0.878)      |              |
| Community characteristics          |              |              |              |              |
| Log (Median household income)      | 0.684 (0.630)| 0.666 (0.629)|            |              |
| % Adults with bachelor’s degree and above | 30.05***     | 30.04***     |            |              |
|                                    | (0.977)      | (0.975)      |            |              |
| % 5–17 year olds in poverty        | −0.137 (1.798)| −0.149 (1.796)|            |              |
| % Household with children and female head | −4.987***    | −4.999***    |            |              |
|                                    | (1.382)      | (1.379)      |            |              |
| % Unemployed                       | −19.66***    | −19.81***    |            |              |
|                                    | (5.320)      | (5.321)      |            |              |
| % Living in same house as last year| 1.736 (1.312)| 1.830 (1.312)|            |              |
| Gini coefficient                   | −0.723 (2.417)| −0.754 (2.418)|            |              |
| City/urban locale                  | −0.0866 (0.270)| −0.129 (0.270)|            |              |
| Suburban locale                    | 0.819*** (0.204)| 0.789*** (0.204)|            |              |
| Town locale                        | 0.431** (0.184)| 0.416** (0.184)|            |              |
| Teacher characteristics            |              |              |              |              |
| % Female teachers                  | 0.616**      | 0.616**      | 1.384**      | 1.384**      |
|                                    | (0.253)      | (0.253)      | (0.506)      | (0.506)      |
| % of teachers with regular state certificate | 0.218       | 0.218       | 1.238**      | 1.238**      |
|                                    | (0.437)      | (0.437)      | (0.552)      | (0.552)      |
| % of teachers with alternative certification | −0.215   | −0.215   |            |              |
|                                    | (0.359)      | (0.359)      |            |              |
| % of teachers with 3–5 years of Experience | 1.238**      | 1.238**      |            |              |
|                                    | (0.552)      | (0.552)      |            |              |
| % of teachers with 6–20 years of Experience | 1.384**      | 1.384**      |            |              |
|                                    | (0.506)      | (0.506)      |            |              |
| % of teachers with 21+ experience  | 0.974*       | 0.974*       | 1.811*       | 1.811*       |
|                                    | (0.555)      | (0.555)      | (0.847)      | (0.847)      |
| Fourth grade                       | 10.63***     | 10.63***     | 10.67***     | 10.67***     |
|                                    | (0.0797)     | (0.0791)     | (0.0791)     | (0.0791)     |
| Fifth grade                        | 21.28***     | 21.32***     | 21.32***     | 21.32***     |
|                                    | (0.0900)     | (0.0885)     | (0.0885)     | (0.0885)     |
| Sixth grade                        | 32.60***     | 32.64***     | 32.64***     | 32.64***     |
|                                    | (0.0977)     | (0.0967)     | (0.0967)     | (0.0967)     |
| Seventh grade                      | 43.41***     | 43.47***     | 43.48***     | 43.48***     |
|                                    | (0.105)      | (0.104)      | (0.104)      | (0.104)      |
| Eighth grade                       | 54.35***     | 54.47***     | 54.47***     | 54.47***     |
|                                    | (0.114)      | (0.112)      | (0.112)      | (0.112)      |
| Constant                           | 104.8***     | 174.3***     | 180.8***     | 181.1***     |
|                                    | (8.196)      | (5.207)      | (8.330)      | (8.472)      |
| Year dummies                       | X            | X            | X            | X            |
| State dummies                      | X            | X            | X            | X            |
| Observations                       | 60,070       | 59,680       | 58,640       | 58,640       |
| $R^2$                              | .015         | .878         | .895         | .895         |

Source. Authors’ calculation based on 2008 to 2009, 2011 to 2012, and 2014 to 2015 Stanford Education Data Archive (SEDA, v. 2.1) combined with 2007 to 2008 and 2011 to 2012 School and Staffing Survey (SASS) and 2015 to 2016 National Teacher and Principal Survey (NTPS).

Note. Errors are clustered within states (presented in parentheses).
*p < .1. **p < .05. ***p < .01. N is rounded to nearest 10.

within the same state do not behave independently of one another, and the estimates from the standard OLS regressions will be biased, and the multilevel models are preferred.

We then conduct the likelihood ratio test, which demonstrates that adding the state-specific intercept into the model improves the fit of the model but the random slope is not necessary. Thus, we treat the effect of teacher salaries similarly for all districts in our teacher-level analysis, and the model estimates a single regression line representing the population average. The district-specific intercept of the model shifts this regression line up or down, maintaining its slope. In our district-level analysis, we consider that all states
have similar salary effects, and the model estimates a single regression line representing the population average with state-specific intercepts.

In Table 3, we present the results for mathematics test scores in panel A and English test scores in panel B, estimated with multilevel models. The math results in columns (1) through (3) in Table 3 are comparable to columns (2) through (4) in Table 1. These coefficients for teacher base salary are almost the same, but the standard errors estimated from the multilevel models are about half of those from the state fixed effects models. If we focus on our preferred model specification, which includes all the control variables, the coefficient for teacher base salary is 1.847 in column (3) of Table 3, which is almost the same as the coefficient of 1.827 from column (4) of Table 1.

The English results in columns (4) to (6) in Table 3 are comparable to columns (2) to (4) in Table 2. The coefficients for teacher base salary are slightly higher in the multilevel models than those in the state fixed effects model, but they are very similar to each other. Again, the standard errors from the multilevel models are much smaller in the multilevel models.

Both approaches rely on different sources of variation, but they produce surprisingly similar results. This suggests that, once we control for key district and community characteristics, the within-state variation is mostly from the unobservable factors that are common among districts within the same state.

The coefficients for higher grades are larger, partly due to the fact that the means of test scores are greater for upper grades. For example, the mean score of the math scores is 229.8 for the 3rd grades and 283 for the eighth grades. To examine if teacher salary is more or less influential for lower or upper grades of students, Table 4 presents the results broken down by student grade for both math in Panel A and English in Panel B. Section 1 reports the results from the state fixed effects model, and section 2 reports the results from the multilevel models. Again, the two methods produce very similar results. Both approaches show that the positive association between teacher salary and test scores is present across different grades, instead of focused on a certain grade, suggesting that better-paid teachers can raise student performance beyond the early childhood education.

To see if a certain race/ethnicity drives this relationship, we estimate the model separately by race/ethnicity of students. Table 5 presents the results for both Math and English. Both models show that White, Black, and Hispanic students perform better in math when teachers earn a higher salary, whereas only Black and Hispanic students’ English scores rise with higher pay. Overall, Hispanic students are the biggest beneficiaries from higher teacher pay. The findings in Table 5 imply that a higher teacher salary is able to reduce the performance gap between both white and black and white and Hispanic students.

We also conduct the separate analyses across different SES of districts to investigate whether the relationship between teacher pay and district performance is generic or distinctive for different SES status. We use a SES index variable, constructed as the first principal component factor score of the measures such as median household income, percent of adults with a bachelor’s degree or higher, single mother-headed household rate, food-stamp receipt rate, poverty rate, and unemployment rate (Fahle et al., 2018). Based on the SES index, we define High-SES group as districts in the top 25% of the SES distribution, Mid-SES group as districts in the middle 50% (25% to 75% of the distribution), and Low-SES group as the bottom 25%.

Table 6 demonstrates the significantly positive relationship between teacher salary and student performance for both High-SES and Mid-SES districts, with a greater magnitude in the High-SES group. A 10% increase in teacher salary is associated, on average, with about a 0.25 points higher math score and a 0.3 points higher English score in the High-SES districts. In the Mid-SES districts, a 10% increase in teacher salary is associated with about 0.15 points higher Math and English scores. For both Math and English, the association between teacher pay and test scores is positive but statistically insignificant in the Low-SES districts. The heterogenous effects by SES are consistent with Han and Maloney (2022) who find significantly positive union effects on test scores in the Mid-SES districts but insignificant union effects in the Low-SES districts.

Our finding that teacher salary has no statistically significant association with test scores in the Low-SES districts suggests that there may exist some threshold of educational environment for teachers to be able to influence student outcome. If so, it may be necessary to equip those disadvantaged districts with sufficient resources and community settings and to combine these efforts with other education initiatives that are complementary to teacher salaries, enabling teacher pay to advance student learning and to address equity concerns.

Districts with strong teachers’ unions tend to pay more to their teachers and increase the usage of other school inputs. For a robustness check, we control for districts’ union density. The alternative results are very similar to those present in Tables 1 through 6, suggesting that higher salary genuinely influences educational outcome, regardless of the strength of teachers’ unions.

To the extent that teacher pay is correlated with teaching experience, our estimates of teacher base salary may be biased. When we drop the district-level variables for teaching experience (i.e. percent of novice teachers and early-career teachers) as a robustness check, however, the alternative results are almost the same as before.

If a district faces greater competition with charter schools and loses its best students to them, the OLS regression will produce biased estimation. Moreover, some charter schools do not follow the pay schedules of their districts; rather, they establish their own pay scheme for individual teachers, which tends to be lower than what the traditional public
Table 3. Estimated Relationship Between Teacher Pay and Districts’ Performance, By Subject Multilevel Mixed Effects Model.

| Variables                                      | Panel A: Mathematics test scores | Panel B: English test scores |
|------------------------------------------------|----------------------------------|------------------------------|
| Log (Base salary)                              | 3.950*** (0.234)                 | 2.045*** (0.217)             |
| District characteristics                        | 1.744*** (0.234)                 | 1.925*** (0.240)             |
| % Hispanic students                             | 1.847*** (0.260)                 | 4.554*** (0.222)             |
| % Black students                                |                                  |                              |
| % Asian students                                |                                  |                              |
| % Native Americans                              |                                  |                              |
| Total enrollment, grades 3–8 (in 1,000)         |                                  |                              |
| % ELL students                                  |                                  |                              |
| % special education students                    |                                  |                              |
| Total teachers (in 1,000)                       |                                  |                              |
| % Free lunch                                   |                                  |                              |
| Community characteristics                       |                                  |                              |
| Log (Median Household Income)                  | 1.655*** (0.297)                 | 0.817*** (0.275)             |
| % Adults with bachelor’s degree and above       | 1.618*** (0.298)                 | 0.789*** (0.275)             |
| % 5–17 year olds in poverty                    | 24.51*** (0.500)                 | 29.90*** (0.461)             |
| % Household with children and female head       | 24.50*** (0.500)                 | 29.90*** (0.461)             |
| % Unemployed                                   | 4.984*** (0.816)                 | −0.503 (0.755)               |
| % Living in same house as last year            | 5.000*** (0.816)                 | −0.516 (0.755)               |
| Gini coefficient                               | −4.652*** (0.664)                | −4.998*** (0.615)            |
| City/urban locale                              | −4.674*** (0.665)                | −5.014*** (0.615)            |
| Suburban locale                                | −19.26*** (2.300)                | −16.34*** (2.140)            |
| Town locale                                    | −19.61*** (2.303)                | −16.56*** (2.141)            |
| Teacher characteristics                        |                                  |                              |
| % Female teachers                              | 0.384*** (0.149)                 | 0.690*** (0.138)             |
| % of teachers with regular state certificate   | −0.0103 (0.265)                  | 0.0850 (0.243)               |
| % of teachers with alternative certification   | −0.0982 (0.221)                  | −0.211 (0.202)               |
| % of teachers with 3–5 years of experience     | 1.096*** (0.319)                 | 1.201*** (0.294)             |
| % of teachers with 6–20 years of experience    | 0.498*** (0.288)                 | 1.335*** (0.265)             |
| % of teachers with 21+ experience              | 0.433 (0.314)                    | 0.881*** (0.289)             |
| Fourth grade                                   | 10.15*** (0.117)                 | 10.67*** (0.107)             |
| Fifth grade                                    | 10.16*** (0.112)                 | 10.67*** (0.107)             |
| Sixth grade                                    | 10.16*** (0.112)                 | 10.67*** (0.107)             |
| Seventh grade                                  | 10.63*** (0.114)                 | 10.67*** (0.107)             |
| Eighth grade                                   | 10.31*** (0.113)                 | 10.67*** (0.107)             |
| Constant                                       | 10.31*** (0.113)                 | 10.67*** (0.107)             |
| Year dummies                                   | 10.31*** (0.113)                 | 10.67*** (0.107)             |
| Observations                                   | 52.09*** (0.124)                 | 54.47*** (0.107)             |
| Number of groups                               | 52.16*** (0.118)                 | 54.47*** (0.107)             |

Source. Authors’ calculation based on 2008 to 2009, 2011 to 2012, and 2014 to 2015 Stanford Education Data Archive (SEDA, v. 2.1) combined with 2007 to 2008 and 2011 to 2012 School and Staffing Survey (SASS) and 2015 to 2016 National Teacher and Principal Survey (NTPS).

Note. Errors are clustered within states (presented in parentheses).

*p < .1. **p < .05. ***p < .01. N is rounded to nearest 10.
### Table 4. Estimated Relationship Between Teacher Pay and Districts’ Performance, By Grade.

| Grade        | Panel A: Mathematics | Panel B: English |
|--------------|----------------------|------------------|
| **Section 1: Results from state fixed effects model** |                       |                  |
| Third grade  | 1.135* (0.582)       | 0.789 (0.692)    |
| Fourth grade | 1.683*** (0.635)     | 1.946*** (0.664) |
| Fifth grade  | 2.341*** (0.628)     | 2.025*** (0.612) |
| Sixth grade  | 0.775 (0.676)        | 1.178* (0.608)   |
| Seventh grade| 2.311*** (0.724)     | 2.493*** (0.592) |
| Eighth grade | 2.847*** (0.837)     | 2.395*** (0.603) |
| **Section 2: Results from multilevel mixed effects model** |                       |                  |
| Third grade  | 1.166** (0.531)      | 0.986 (0.622)    |
| Fourth grade | 1.657*** (0.561)     | 2.072*** (0.599) |
| Fifth grade  | 2.248*** (0.591)     | 2.124*** (0.567) |
| Sixth grade  | 0.674 (0.623)        | 1.254** (0.563)  |
| Seventh grade| 2.170*** (0.656)     | 2.586*** (0.542) |
| Eighth grade | 2.677*** (0.748)     | 2.431*** (0.544) |

*Source: Authors’ calculation based on 2008 to 2009, 2011 to 2012, and 2014 to 2015 Stanford Education Data Archive (SEDA, v. 2.1) combined with 2007 to 2008 and 2011 to 2012 School and Staffing Survey (SASS) and 2015 to 2016 National Teacher and Principal Survey (NTPS). Note. Errors are clustered within states (presented in parentheses). *p < .1. **p < .05. ***p < .01. All models include the full set of controls listed in Tables 1 and 2.

### Table 5. Estimated Relationship Between Teacher Pay and Districts’ Performance, By Students’ Race/Ethnicity.

| Race/ethnicity | Panel A: Mathematics | Panel B: English |
|----------------|----------------------|------------------|
| **Section 1: Results from state fixed effects model** |                       |                  |
| White         | 1.161** (0.521)      | 0.360 (0.522)    |
| Black         | 1.687** (0.858)      | 2.512*** (0.918) |
| Hispanic      | 2.338*** (0.782)     | 2.932*** (0.834) |
| Asian         | 1.321 (1.793)        | 1.474 (1.764)    |
| **Section 2: Results from multilevel mixed effects model** |                       |                  |
| White         | 1.167*** (0.281)     | 0.380 (0.267)    |
| Black         | 1.658*** (0.492)     | 2.569*** (0.496) |
| Hispanic      | 2.257*** (0.436)     | 2.956*** (0.447) |
| Asian         | 1.284 (0.949)        | 1.466 (0.892)    |

*Source: Authors’ calculation based on 2008 to 2009, 2011 to 2012, and 2014 to 2015 Stanford Education Data Archive (SEDA, v. 2.1) combined with 2007 to 2008 and 2011 to 2012 School and Staffing Survey (SASS) and 2015 to 2016 National Teacher and Principal Survey (NTPS). Note. Errors are clustered within states (presented in parentheses). *p < .1. **p < .05. ***p < .01. All models include the full set of controls listed in Tables 1 and 2.

### Table 6. Estimated Relationship Between Teacher Pay and Districts’ Performance, By Districts SES Status.

| SES status | Panel A: Mathematics | Panel B: English |
|------------|----------------------|------------------|
| **Section 1: Results from state fixed effects model** |                       |                  |
| High-SES districts | 2.465*** (0.916)     | 2.995*** (0.802) |
| Mid-SES districts | 1.455*** (0.667)     | 1.453*** (0.648) |
| Low-SES districts | 0.613 (0.988)        | 0.347 (0.955)    |
| **Section 2: Results from multilevel mixed effects model** |                       |                  |
| High-SES districts | 2.386*** (0.491)     | 2.985*** (0.443) |
| Mid-SES districts | 1.436*** (0.369)     | 1.500*** (0.348) |
| Low-SES districts | 0.394 (0.523)        | 0.343 (0.494)    |

*Source: Authors’ calculation based on 2008 to 2009, 2011 to 2012, and 2014 to 2015 Stanford Education Data Archive (SEDA, v. 2.1) combined with 2007 to 2008 and 2011 to 2012 School and Staffing Survey (SASS) and 2015 to 2016 National Teacher and Principal Survey (NTPS). Note. Errors are clustered within states (presented in parentheses). *p < .1. **p < .05. ***p < .01. All models include the full set of controls listed in Tables 1 and 2.
schools offer. To address these issues, we add the percent of students in charter schools as an additional control variable. The alterative results remain almost the same.

Our multilevel models do not include state dummies, so one may be concerned with the potential bias due to between-state variation. Because our analysis includes numerous covariates at the district level, most of the between-state variation is likely to be captured by those variables. Nonetheless, as a final robustness check, we also control for two state-level variables that we bring from the data provided by the U.S. Bureau of Economic Analysis (BEA): state population from years 2009, 2012, and 2015 and the GDP growth rate from years 2008 to 2009, 2011 to 2012, and 2014 to 2015. The alternative results after controlling for these two variables are almost identical to those presented in Table 3, and the coefficients for these variables are not statistically significant.

**Discussion**

One of the key findings of our study is that a higher teacher salary is associated with a reduction in educational inequality between white and black, as well as white and Hispanic students, because the estimated relationships are stronger for minority students, except for Asian students. Considering the close link between educational inequality and income inequality (Chetty et al., 2014), our study suggests that raising teacher salary is linked to a decreased performance gap across different racial groups of students and higher intergenerational mobility for minority families in the long run.

Districts that provide higher teacher compensation are better able to attract high-quality teachers and retain them in schools. Moreover, teachers’ morale and enthusiasm are more likely to be higher when they are paid more. Echoing the benefits of higher teacher pay, advocates for the pay-for-performance system urge an increase in the compensation of teachers based on their productivity, measured by educational outcomes of their students. However, our study finds that universal and unconditional increase in teacher base salary (i.e., across-the-board base salary increases), regardless of student performance or any supplementary compensation tied with educational outcomes, is associated with improved districts’ academic performance. In particular, if districts are to adopt a merit pay system at the expense of reducing base salaries to balance the budget, we predict, based on our results, that their educational outcomes may suffer, rather than improve.

Our estimates on salary effects concern all students (nationally representative sample of all school districts). The gain is relatively small (between 1 and 3 test score points, depending on the racial/ethnic group and grade), but it applies to the general student population, potentially producing a large accumulated gain over the entire student body in the country. The coefficients are always positive and statistically significant for most subgroups, and no subgroup shows negative salary effects. Moreover, the results in this study may be a lower bound estimate for the influence of teacher salaries on performance because our samples are incumbent teachers (i.e., a static sample). If the new policy intervention is to raise teacher salary significantly, we are not taking into account any potential changes in the composition of teachers that may occur driven by new teachers coming into the profession and the relative credentials of the teaching workforce over time. With a significantly higher teacher salaries, according to the channels reviewed above, schools would be able to retain and attract highly qualified teachers, and help build a relatively stronger teaching workforce. Eventually, the long-run effect of higher base salaries on district performance may be larger than the short-run effects that we estimate in this study.

Our findings suggest that a substantial increase in teacher pay may be needed for a significant and sustained improvement in student performance. This is aligned with the argument by Temin (2002) that a small increase in teacher pay may not result in a meaningful improvement. By presenting the multiple equilibria of the US teachers’ labor market based on Akerlof’s “Lemon” model, Temin argues that the US cannot escape from an inferior equilibrium, where lower teacher pay is matched with teachers with low productivity, unless districts pay substantially more to teachers.

It is noteworthy that there exists a large variance in teacher base salary within each state, and the within-state variance greatly differs by state. Overall, the variance in teacher base salary tends to be larger in pro-union states than in anti-union states. For instance, the variance is about $12,000 in California, which is approximately 25% of its average teacher base salary. On the other hand, the variance is about $5,000 in South Carolina, and it is around 10% of its average teacher base salary. Thus, a 10% increase in the teacher base salary may be considered as substantial in South Carolina but not in California.

Most analyses of student performance and school effectiveness in elementary schools find that household and neighborhood conditions outweigh the effects of specific school characteristics. Indeed, the positive association between teacher salary and district performance we have identified in this study is modest in magnitude. Nevertheless, our study finds that teacher salary significantly correlates with increased English test scores, which is harder to raise than math test scores, demonstrating that changing the educational inputs in a proper way can improve student outcomes, and educators and policy makers should continue to search for the most effective intervention for that goal.

**Conclusion**

This research draws on evidence from a rich data set linking district performance on achievement test scores to average teacher base salary nationwide. We employ two
identification strategies: state fixed effects and multilevel mixed effects linear models.

Our findings consistently show a significantly positive association between teacher base salary and districts’ performance. We find that both mathematics and English test scores are significantly higher in districts that offer a higher base salary to teachers, compared to those in districts with a lower teacher base salary. In both state fixed effects and multilevel models, we find that a 10% increase in teacher salary is associated with an increase of about 0.2 points in test scores in both subjects. The association between teacher salaries and performance is more intense in higher grade-levels than in lower grade-levels. Overall, these findings shed light both on how improving teacher pay directly correlates with student performance and on existing debates on policies seeking teacher compensation reforms.

We also find that higher teacher salaries are associated with the reduced achievement gap between white and black, and between white and Hispanic students, because the coefficient is greater for minority students (except for Asian students). There exists a significantly positive relationship between teacher salary and student performance for the districts with high- and medium-level socioeconomic status, but not in the districts with low socioeconomic status.

Even with the state fixed effects and multi-level models, we are not able to make a causal claim of our findings, because we cannot fully account for all of the confounding factors (both observable and unobservable) at the district level due to the limitation of cross-sectional data. One such factor is districts’ general political attitude toward public education. Districts that are more likely to adopt public policies will allocate more resources to public education, raising teacher pay while influencing student performance. Based on the similar results from two very different methods, however, our estimates appear to offer strong evidence on the positive relationship between teacher salaries and district performance.

Our study utilizes standardized test scores to assess student performance in districts. It is noteworthy that literature is not in consensus regarding whether the test score is the single most important measurement for educational outcomes, and there may be other metrics that can better capture educational success. Therefore, exploring different measures of student outcomes is an important task in examining the effect of teacher compensation on public education. This subject is left for future study.

Acknowledgments
We thank the National Center for Educational Statistics (NCES) for kindly providing us with the data. We also thank the National Bureau of Economic Research (NBER) and the Economic Policy Institute (EPI) for providing us with the necessary facilities and assistance. We appreciate John Schmitt’s comments on an earlier version of this manuscript.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs
Emma García https://orcid.org/0000-0001-7255-1651
Eunice S. Han https://orcid.org/0000-0002-0475-0488

Supplemental Material
Supplemental material for this article is available online.

Notes
1. For recent evidence on these achievement gaps and their trends, see Hanushek et al. (2019) and Reardon and Portilla (2016) among others.
2. Pay is not the only obstacle for attracting students into teaching. They also cite lack of opportunities for career development and others as reasons why teaching is unattractive, which is consistent with other empirical research more broadly.
3. Growing evidence on this front also exists from papers examining other forms of wages, such as salary schedules or incentives programs. On teacher salary schedules, research by Murnane and Olsen (1990) finds that an increase of $1,000 in each salary step would increase teachers’ mean duration in that district by 2 to 3 years (data from North Carolina). On teacher pay for performance strategies, findings are mixed and driven by multiple parameters—including assumptions about its components and functioning, amount offered, starting quality level, country where the strategy is implemented, etc. (de Ree et al., 2016; Figlio & Kenny, 2007; Ganimian & Murnane, 2016; Gius, 2012; Imberman, 2015; Jones, 2013). For example, Figlio and Kenny (2007) find a positive association between student achievement and teacher incentives, but the authors share the caveat that they cannot rule out if incentives lead to more effort, and this leads to higher scores, or if there is any district quality component by which some districts are more likely to adopt these incentives.
4. See a discussion of the teacher wage penalty trends in Allegretto and Mishel (2019) and of their values internationally in Hanushek et al., (2019).
5. In 2008 to 2009, 97% of beginning teachers whose first-year base salary was $40,000 or more were teaching, versus 87% of those whose first-year base salary was less than $40,000; in 2011 to 2012, these shares (using the same base salary cutoff) were 89% versus 80% (based on the Beginning Teacher Longitudinal Study, for the United States).
6. Note that to a great extent, the estimates rely on student performance measured by standardized tests, but other measures
of student performance are sometimes also considered. These include continuation in school, dropout behavior, and subsequent labor market earnings” (Hanushek 2003, F75).

7. On the one hand, Britton and Propper (2016) acknowledge Hanushek’s assessment of the mixed evidence of the impact of teacher pay on school performance, but note that “but more recent evidence suggests otherwise.” On the other hand, Hanushek et al. (2019) continue to argue that “studies conducted within specific countries tend to find that salaries are unrelated to effectiveness,” but that there is limited cross-country evidence suggesting that “students perform better where teachers are better paid” in recent studies. Heavier reliance on alternative methods beyond meta-analyses that built on correlational studies and country-level analyses should be expected to shed further light on these relationships.

8. A criticism of Loeb and Page (2000) is that the relationship is sensitive to controls for alternative earnings opportunities and other factors that affect both teacher labor market decisions and student achievement (Hanushek & Rivkin, 2007, p. 80).

9. One criticism to their analysis, performed at the state level, is that both aggregation at the state level and omitted variables may bias their results. See Loeb and Page (2000) for a discussion.

10. To obtain the district-level teacher base salaries, we aggregated the teacher-level responses to the district level, weighted by teachers’ final sample weights.

11. In all, about 12,000 districts are represented in the SEDA, but the extent of coverage in the SEDA data varies across year, subject, and subgroup, for several reasons. In some cases, a variety of assessment tests were used within state or within district, precluding the construction of usable means and standard deviations. Low rates of student participation, incomplete reporting, and small cell sizes for subgroups also altered coverage (see Fahle et al., 2018).

12. Our base salary information is from the teacher-level questionnaire asking “During the current school year, what is your academic year base teaching salary?” This information is available in both the SASS and the NTPS. The base teaching salary does not include any other compensation such as additional compensation from this school system for extracurricular (such as coaching, student activity sponsorship, or teaching evening classes) or additional activities or other income from this school system (such as a merit pay bonus or state supplement).

13. Because the NAEP tests are only administered in the fourth and eighth grades, the test scores for other grades included in the SEDA are imputed values, based on these NAEP test scores. See Fahle et al. (2017) for details.

14. We test for multicollinearity among these covariates, and the degree of correlation is small enough so that they do not pose any problem in our models. All of the variance inflation factors (VIF), which indicate the extent to which multicollinearity is present in a regression analysis, among the covariates are well below the rule of thumb of 10, and most of the IVF are less than 5.

15. The SASS is not designed to construct panel data, and majority of districts only show up in one wave of the survey. Thus, we do not employ district fixed effects.

16. Generally speaking, the coefficients are almost identical, and standard errors are somewhat higher under the meta-analytic regressions. For instance, the coefficient and SE for log(base salary) change from 1.827*** (0.478) to 1.783*** (0.653) in model (4) of Table 1, and they change from 1.808*** (0.453) to 1.819*** (0.747) in model (4) of Table 2.

17. Intra-class correlation $\hat{\rho}$ is a summary of the proportion of the outcome variation that is attributable to the differences across entities, and it is calculated as $\hat{\rho} = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \hat{\sigma}_k^2}$, ranging between 0 and 1.

References

Adamson, F., & Darling-Hammond, L. (2012). Funding disparities and the inequitable distribution of teachers: Evaluating sources and solutions. Education Policy Analysis Archives, 20, 37.

Allegretto, S., & Mishel, M. (2019). The teacher weekly wage penalty hit 21.4 percent in 2018, a record high: Trends in the teacher wage and compensation penalties through 2018. Economic Policy Institute and the Center on Wage & Employment Dynamics at the University of California, Berkeley, April 2019. https://www.epi.org/publication/the-teacher-weekly-wage-penalty-hit-21-4-percent-in-2018-a-record-high-trends-in-the-teacher-wage-and-compensation-penalties-through-2018/

Boarini, R., & Luédmann, E. (2009). The role of teacher compensation and selected accountability policies for learning outcomes: An empirical analysis for OECD countries. OECD Journal: Economic Studies, 2009(1), 1–21. https://doi.org/10.1787/eco_studies-v2009-art10-en

Britton, J., & Propper, C. (2016). Teacher pay and school productivity: Exploiting wage regulation. Journal of Public Economics, 133, 75–89. https://doi.org/10.1016/j.jpubeco.2015.12.004

Card, D., & Krueger, A. B. (1992). Does school quality matter? Returns to education and the characteristics of public schools in the United States. Journal of Political Economy, 100(1), 1–40.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. American Economic Review, 104(9), 2633–2679.

Clotfelter, C., Ladd, H. F., Vigdor, J., & Wheeler, J. (2006). High-poverty schools and the distribution of teachers and principals. North Carolina Law Review, 85, 1345.

Council of Economic Advisers. (2016). Labor market monopsony: Trends, consequences, and policy responses. Issue Brief, October 2016. https://obamawhitehouse.archives.gov/sites/default/files/page/files/20161025_monopsony_labor_mrk_ cee.pdf

Croft, M., Guffy, G., & Vitale, D. (2018). Encouraging more high school students to consider teaching. ACT Policy Research. https://www.act.org/content/dam/act/unsecured/documents/pdfs/Encouraging-More-HS-Students-to-Consider-Teaching.pdf

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. Education Policy Analysis Archives, 8, 1–44. https://doi.org/10.14507/epaa.v8n1.2000

de Ree, J., Muralidharan, K., Pradhan, M., & Rogers, H. (2016). Double for nothing? Experimental evidence on the impact of an unconditional teacher salary increase on student performance in Indonesia (NBER Working Paper 21806). NBER.
Dolton, P., & Marcenaro-Gutiérrez, O. D. (2011). If you pay peanuts do you get monkeys? A cross-country analysis of teacher pay and pupil performance. *Economic Policy*, 26(1), 5–55. https://doi.org/10.10111/j.1468-0327.2010.00257.x

Fahle, E. M., DiSalvo, R., Ho, A. D., Kalogrides, D., Reardon, S. F., & Shear, B. R. (2017). *Stanford education data archive (Version 2.0)*. http://purl.stanford.edu/db586ns4974

Fahle, E. M., Shear, B. R., Kalogrides, D., Reardon, S. F., DiSalvo, R., & Ho, A. D. (2018). *Stanford education data archive: Technical documentation (Version 2.1)*. http://purl.stanford.edu/db586ns4974.

Figlio, D. N. (1997). Teacher salaries and teacher quality. *Economic Letters*, 55(2), 267–271. https://doi.org/10.1016/S0165-1765(97)00070-0

Figlio, D. N., & Kenny, L. W. (2007). Individual teacher incentives and student performance. *Journal of Public Economics*, 91(5), 901–914. https://doi.org/10.1016/j.jpubeco.2006.10.001

Ganigian, A. J., & Murnane, R. J. (2016). Improving education in developing countries: Lessons from rigorous impact evaluations. *Review of Educational Research*, 86(3), 719–755. https://doi.org/10.3102/0034654316527499

Gius, M. (2012). The effects of merit pay on teacher job satisfaction. *Applied Economics*, 45(31), 4443–4451. https://doi.org/10.1080/00036846.2013.788783

Garcia, E., & Weiss, E. (2019). U.S. schools struggle to hire and retain teachers: The second report in ‘The Perfect Storm in the Teacher Labor Market’ series. Economic Policy Institute.

Glewwe, P., Hanushek, E. A., Humpage, S., & Ravina, R. (2014). School resources and educational outcomes in developing countries: A review of the literature in developing countries from 1990 to 2010. In P. Glewwe (ed.), *Education policy in developing countries*. University of Chicago Press.

Goldhaber, D., DeArmond, M., Player, D., & Choi, H. (2008). Why do so few public school districts use merit pay? *Journal of Education Finance*, 33(3), 262–289. https://www.jstor.org/stable/40704329

Goldhaber, D., Lavery, L., & Theobald, R. (2014). Uneven playing field? Assessing the inequity of teacher characteristics and measured performance across students (CEDR Working Paper 2014-4). Center for Education Data and Research, University of Washington.

Gray, L., & Taie, S. (2015). Public school teacher attrition and mobility in the first five years: Results from the first through fifth waves of the 2007-08 beginning teacher longitudinal study: First look. National Center for Education Statistics, U.S. Department of Education. https://nces.ed.gov/pubs2015/2015337.pdf

Grissom, J. A., Viano, S. L., & Selin, J. L. (2015). Understanding employee turnover in the public sector: Insights from research on teacher mobility. *Public Administration Review*, 76(2), 241–251. https://doi.org/10.1111/puar.12435

Han, E. S., & Maloney, T. (2022). Teachers unionization, socioeconomic status, and student performance in the US. *American Journal of Education. Advance online publication*. https://doi.org/10.1086/717673

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Education Policy Analysis Archives*, 19(2), 141–161. https://doi.org/10.3102/01623737019002141

Hanushek, E. A. (2003). The failure of input-based schooling policies. *Economic Journal*, 113 (485), F64-F98. https://doi.org/10.1111/1468-0297.00999

Hanushek, E. A. (2015). *International encyclopedia of the social & behavioral sciences* (2nd ed., Vol. 7). Hoover Institution, Stanford University. https://doi.org/10.1016/B978-0-08-097086-8.92052-X

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (1998). Teachers, schools, and academic achievement (NBER Working Paper No. W6961). NBER.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (1999). Do higher salaries buy better teachers? (NBER Working Paper No. 7082). NBER.

Hanushek, E. A., Peterson, P. E., Talpey, L. M., & Woessmann, L. (2019). The unwavering SES achievement gap: Trends in U.S. student performance (NBER Working Paper No. 25648). NBER.

Hanushek, E. A., Piopiunik, M., & Wiederhold, S. (2019). Do smarter teachers make smarter students? *Education Next*, 2019, 57–64. https://www.educationnext.org/do-smarter-teachers-make-smarter-students-international-evidence-cognitive-skills-performance/

Hanushek, E. A., & Rivkin, S. G. (2007). Pay, working conditions, and teacher quality. *The Future of Children*, 17(1), 69–86. https://doi.org/10.1353/foc.2007.0002

Hanushek, E. A., Rivkin, S. G., & Shimao, J. C. (2016). Dynamic effects of teacher turnover on the quality of instruction (Working Paper No. 170). National Center for Analysis of Longitudinal Data in Education Research (CALDER). November 2016. https://caldercenter.org/publications/dynamic-effects-teacher-turnover-quality-instruction

Hartel, E. H. (2013). Reliability and validity of inferences about teachers based on student test scores, the 14th William H. Angoff memorial lecture. ETS. https://www.ets.org/Media/Research/pdf/PICANG14.pdf

Hendricks, M. D. (2014). Does it pay to pay teachers more? Evidence from Texas. *Journal of Public Economics*, 109, 50–63. https://doi.org/10.1016/j.jpubeco.2013.11.001

Hendricks, M. D. (2016). Do higher salaries differentially retain high-ability teachers? https://ssrn.com/abstract=2822042

Imberman, S. A. (2015). How effective are financial incentives for teachers? *IZA World of Labor, Institute of Labor Economics (IZA)*, 158. https://doi.org/10.15185/izawol.158

Isenberg, E., Max, J., Gleason, P., Potamites, L., Santillano, R., Hock, H., & Hansen, M. (2013). Access to effective teaching for disadvantaged students. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Jackson, C. K. (2018). Does school spending matter? The new literature on an old question (NBER Working Paper No. 25368). NBER.

Jackson, C. K., & Bruegmann, E. (2009). Teaching students and teaching each other: The importance of peer learning for teachers. *American Economic Journal: Applied Economics*, 1(4), 85–108. https://doi.org/10.1257/app.1.4.85

Jones, M. D. (2013). Teacher behavior under performance pay incentives. *Economics of Education Review*, 37, 148–164. https://doi.org/10.1016/j.econedurev.2013.09.005

Katz, L. F. (1986). Efficiency wage theories: A partial evaluation (NBER Working Paper No. 1906). NBER.
Katz, V. (2018). Teacher retention: Evidence to inform policy. Policy Brief, Curry School and Batten School, EdPolicyWorks, University of Virginia. https://curry.virginia.edu/sites/default/files/uploads/epw/Teacher%20Retention%20Policy%20Brief.pdf

Kraft, M. A., & Papay, J. P. (2014). Can professional environments in schools promote teacher development? Explaining heterogeneity in returns to teaching experience. Educational Evaluation and Policy Analysis, 36(4), 476–500. https://doi.org/10.3102/0162373713519496

Krueger, A. B., & Summers, L. H. (1988). Efficiency wages: Models of unemployment, inter-industry wage structure, Econometrica, 56(2), 259–293. https://doi.org/10.2307/1911072

Ladd, H. F., & Sorensen, L. C. (2016). Returns to teacher experience: Student achievement and motivation in middle school. Education Finance and Policy, 12(2), 241–279. https://doi.org/10.1162/EDFP_a_00194

Lafortune, J., Rothstein, J., & Schanzenbach, D. W. (2018). School finance reform and the distribution of student achievement. American Economic Journal: Applied Economics, 10(2), 1–26. https://doi.org/10.1257/app.20160567

Leigh, A. (2012). Teacher pay and teacher aptitude. Economics of Education Review, 1, 41–53. https://doi.org/10.1016/j.econedurev.2012.02.001

Lin, T.-C. (2010) Teacher salaries and student achievement: The case of Pennsylvania. Applied Economics Letters, 17(6), 547–550. https://doi.org/10.1080/13504850802167223

Loeb, S., Darling-Hammond, L., & Luczak, J. (2005). How teaching conditions predict teacher turnover in California schools. Peabody Journal of Education, 80(3), 44–70. https://doi.org/10.1207/s15327930pe8003_4

Loeb, S., & Page, M. E. (2000). Examining the link between teacher wages and student outcomes: The importance of alternative labor market opportunities and nonpecuniary variation. The Review of Economics and Statistics, 82(3), 393–408. https://doi.org/10.1162/003465300558894

Manski, C. F. (1987). Academic ability, earnings, and the decision to become a teacher: Evidence from the national longitudinal study of the high school class of 1972. In D. A. Wise (Ed.), Public sector payrolls (pp. 291–316). University of Chicago Press.

Marchand, J., & Weber, J. G. (2019). How local economic conditions affect school finances, teacher quality, and student achievement: Evidence from the Texas Shale Boom. Journal of Policy Analysis and Management, 0(0), 1–28. https://doi.org/10.1002/pam.22171

Murnane, R. J., & Cohen, D. K. (1986). Merit pay and the evaluation problem: Why most merit pay plans fail and a few survive. Harvard Educational Review, 56(1), 1–17. https://doi.org/10.17763/haer.56.1.8q2334243271116

Murnane, R. J., & Olsen, R. J. (1989). The effect of salaries and opportunity costs on duration in teaching: Evidence from Michigan. The Review of Economics and Statistics, 71(2), 347–352. https://doi.org/10.2307/1926983

Murnane, R. J., & Olsen, R. J. (1990). The effects of salaries and opportunity costs on length of stay in teaching: Evidence from North Carolina. The Journal of Human Resources, 25(1), 106–124. https://doi.org/10.2307/145729.

Murnane, R. J., Willett, J. B., & Levy, F. (1995). The growing importance of cognitive skills in wage determination. The Review of Economics and Statistics, 77(2), 251–266. https://doi.org/10.2307/2109863

Nagler, M., Piopiunik, M., & West, M. R. (2017). Weak markets, strong teachers: Recession at career start and teacher effectiveness (Working Paper No. 21393). NBER. http://www.nber.org/papers/w21393

National Center for Education Statistics (NCES) U.S. Department of Education. (2007–2008). Licensed microdata from the 2007–2008 Schools and Staffing Survey (SASS). National Center for Education Statistics (U.S. Department of Education).

National Center for Education Statistics (NCES) U.S. Department of Education. (2011–2012). Licensed microdata from the 2011–2012 Schools and Staffing Survey (SASS). National Center for Education Statistics (U.S. Department of Education).

Podolsky, A., Kini, T., Darling-Hammond, L., & Bishop, J. (2019). Strategies for attracting and retaining educators: What does the evidence say? Education Policy Analysis Archives, 27, 38. https://doi.org/10.14507/epaa.27.3722

Reardon, S. F., Kallogrides, D., & Shores, K. (2019) The geography of racial/ethnic test score gaps. American Journal of Sociology, 124(4), 1164–1221.

Reardon, S. F., & Portilla, X. A. (2016). Recent trends in income, racial, and ethnic school readiness gaps at kindergarten entry. AERA Open, 2(3), 1–18. https://doi.org/10.1177/2332858416657343

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. American Educational Research Journal, 50, 4–36. https://doi.org/10.3102/0021935412463813

Sass, T. R., Hannaway, J., Xu, Z., Figlio, D. N., & Feng, L. (2012). Value added of teachers in high-poverty schools and lower poverty schools. Journal of Urban Economics, 72(2), 104–122.

Sorensen, L. C., & Ladd, H. (2018). The hidden costs of teacher turnover (Working paper no. 203-0918-1). National Center for Analysis of Longitudinal Data in Education Research (CALDER). https://caldercenter.org/publications/hidden-costs-teacher-turnover

Stiglitz, J. (1986). Theories of wage rigidities. In J. Butkiewicz, K. J. Koford, & J. B. Miller (Eds.), Keynes’ economic legacy: Contemporary economic theories (pp. 153–206). Praeger Publishers.

Stockard, J., & Lehman, M. B. (2004). Influences on the satisfaction and retention of 1st-year teachers: The importance of effective school management. Educational Administration Quarterly, 40(5), 742–771. https://doi.org/10.1177/0013161X04268844

Temen, T. (2002). Teacher quality and the future of America. Eastern Economic Journal, 28(3), 285–300.

Weiss, A. (1980). Job queues and layoffs in labor markets with flexible wages. Journal of Political Economy, 88(3), 526–538. https://doi.org/10.1086/260884

Weiss, A. (2017). Efficiency wages: Models of unemployment, layoffs, and wage dispersion. Princeton University Press.