11-5-2020

Contextualizing Educational Disparities and the Evaluation of Teacher Quality

Dara Shifrer
Portland State University, dshifrer@pdx.edu

Citation Details
Shifrer, D. (2020). Contextualizing Educational Disparities and the Evaluation of Teacher Quality. Social Problems. (Accepted Manuscript)

Contextualizing Educational Disparities and the Evaluation of Teacher Quality

Dara Shifrer
Department of Sociology
Portland State University

Abstract: Value added scores, statistical estimates of teacher quality, are exemplar of neoliberal logic. The higher average scores of teachers of socially advantaged students raise concerns that scores are inaccurate and unfair, and propagate decontextualized neoliberal understandings of the nature of learning and teachers' work. This study uses longitudinal data on around 4,500 teachers in a large urban district between 2007-08 through 2012-13 to follow individual teachers as they switch into schools of different "performance levels" over time. Fixed-intercept models tracking individual teachers between 2007-08 and 2012-13 showed scores increased for teachers who switched into high-performing schools, and decreased for teachers who switched into low-performing schools. Particularly indicative of scores biased by contextual factors outside teachers' control, score changes for mobile teachers are partially attributable to shifts in the economic status and race of students in teachers' classrooms and schools. Understanding how neoliberalism operates within education provides sociological insight into how neoliberalism is legitimated and perpetuated in other central social institutions, like the criminal justice system, the environment, gender, sexuality, and health.

Keywords: neoliberalism; accountability; structural inequality; educational disparities; educational policy; teachers

Published online in *Social Problems* Advance Articles (https://doi.org/10.1093/socpro/spaa044). This work was supported by the Houston Education Research Consortium (HERC). HERC, directed by Dr. Ruth López Turley, was funded by the Laura and John Arnold Foundation and the Houston Endowment Inc. This project also benefitted from the assistance of district research staff, HERC colleagues, anonymous reviewers, and Drs. Sarah Blanchard, Thurston Domina, Shani Evans, Rachel Fish, Angela Frederick, Eric Grodsky, Laura Mauldin, Anna Mueller, Jennifer Pearson, Heather Price, Carrie Shandra, and anonymous reviewers.

Contextualizing Educational Disparities and the Evaluation of Teacher Quality

With the proliferation of educational data, identifying effective teachers should be easier than ever. Value-added scores are statistical estimates of teacher quality based on students' test score gains (Harris and Herrington 2015). Although these scores ostensibly account for factors outside of teachers' control that influence student achievement gains (McCaffrey et al. 2003), teachers in schools serving socially advantaged students typically have higher scores (Rice 2010; Sass et al. 2012). Because of the long line of sociological research showing that achievement is in larger part a function of student social background than schools or teachers (Coleman et al. 1966; Gamoran and Long 2006; Hill 2016; Quinn 2015), this raises concern that value-added methodologies still provide inaccurate and unfair measures of teacher effects. This study uses longitudinal data on around 4,500 teachers in a large urban district between 2007-08 through 2012-13 to investigate the degree to which value-added scores reflect the qualities of teachers, or the qualities of their students and schools.

Because value-added scores are defined and applied as stable reliable estimates of individual teacher quality (McCaffrey et al. 2003), they should not be contextually variable, that is, should not vary on the basis of changes in the characteristics of their students and schools. In this study, fixed-intercept models longitudinally track the value-added scores of individual teachers as they switch into schools of different performance levels. To tangibly link to how schools are evaluated and perceived by policymakers and the public (Finster and Miller 2014), school "performance level" is measured by averaging students' test scores. Finally, a decomposition technique explores the degree to which the estimated effect of switching school-performance-levels on value-added scores is mediated by changes in the characteristics of teachers' students and schools. Morgan and Shackelford (2018) specifically call for the study of

value-added scores to increase sociological understanding of the effects of schools and teachers. This study employs scores from the Education Value-Added Assessment System (EVAAS), the most widely implemented value-added methodology in the US (Amrein-Beardsley and Collins 2012; Kupermintz 2003). Despite its prevalence, EVAAS is largely absent from the previous literature evaluating value-added methodologies, in part because the EVAAS approach is proprietary and protected (Amrein-Beardsley and Collins 2012; McCaffrey and Hamilton 2007; Vosters, Guarino and Wooldridge 2018). By not basing analyses specifically on the EVAAS methodology, this study creatively facilitates consideration of a scoring method central in dramatic shifts in how we understand teacher effects and the organization of schools.

Neoliberalism—marked by a value for free markets, the privatization of public goods, and limited state intervention—proliferated globally, but particularly in the US, in the mid-20th century (Trumpy 2008). Neoliberal logic privileges individual accountability over structure and context, and emphasizes performativity, productivity, and efficiency (Ball 2016). Among all the examples of neoliberalism in contemporary education systems (e.g., school accountability systems, standardized tests), value-added methods are exemplar. These methods prioritize individualistic understandings of teachers' work over contextual understandings, exact teacher accountability through surveillance and quantification, and are sometimes explicitly used to inspire competition among teachers (Garver 2019; Hutt and Tang 2013). The implications of these findings are important given the proliferation of data and neoliberal logics in multiple realms, including criminal justice (Goodman 2012), the environment (Vasseur 2016), gender and sexuality (Valocchi 2017), health (Nancarrow and Borthwick 2005), and parenting (Reich 2014). Understanding how neoliberal logics operate within a central social institution like education provides a parallel for sociological investigations into neoliberalism in other social institutions.

Tied to foundational philosophies of how the world works, the expansion of neoliberalism parallels rising tides of conservatism in the US and Europe, increasing inequality, and dwindling structural supports (Piketty and Saez 2004). Sociologists already understand the centrality of inequality outside of schools for the work of teachers and for students' outcomes. This study contributes an examination of how neoliberal tools perpetuate perverted understandings of teachers' work and of the process of learning, to the detriment of policy and advancements in sociological knowledge.

TEACHER EFFECTS AND NEOLIBERALISM

Teacher effects and the organization of schools has a rich sociological history. Just over fifty years ago, the seminal 'Coleman Report' exposed differences across homes rather than schools and teachers as the primary source of differences in achievement (Coleman et al. 1966). Although decades of research confirm these findings with different data and more sophisticated methodology, neoliberally-informed educational policy continues to emphasize teachers (Loveless 2003). Around the same time, researchers described schools as bureaucratic in some ways, and not bureaucratic in other ways, particularly with the loose articulation between subunits (Bidwell 1965). The "loosely coupled" nature of schools actually increases professional expectations for teachers, ostensibly leaving them masters of their own classrooms (Bidwell 1965; Weick 1976). Neoliberalism is marked by an emphasis on individual accountability over the contributions of structure and context (Hill Collins and Bilge 2016). Whereas Coleman conceptualized teachers as a school characteristic (Morgan and Shackelford 2018), organizational conceptualizations of schools laid the foundation for a neoliberal vision of teacher's work, with student achievement depicted as the result of teachers' individualized efforts within the classroom.

Loose coupling facilitates local control, yet our contemporary education system is marked by competing emphases on both local control and standardization, i.e., decentralization and centralization (Renzulli 2014). Local control is argued to facilitate student performance but then criticized for masking failures to produce achievement (Meyer and Rowan 2006). With standardization through bureaucratic principles and federal control emphasized as a means of ensuring success for all (Renzulli 2014), neoliberal logics like accountability, performativity, productivity, and efficiency flourish (Ball 2016). In 2001, No Child Left Behind (NCLB) federally mandated states implement accountability systems with tests aligned with learning standards and accountability ratings for districts and schools (Morgan and Shackelford 2018). Just as neoliberalism has more broadly shifted the burden of economic risk from government to individuals and families since the 1970s (Silva 2019), NCLB shifted the burden for educational outcomes from broad structural change to teachers and schools (Schneider, Grogan and Maier 2011). Schools are now ranked as high- or low-performing on the basis of standardized test scores (Ambrosio 2013), which facilitates neoliberal priorities for endogenous privatization (introducing choice and competition) and exogenous privatization (the introduction of new educational providers) (Ball 2016). The Every Student Succeeds Act replaced NCLB In 2015 but also emphasizes standards and accountability. Prioritizing individualistic understandings of teachers' work, exacting teacher accountability through teacher surveillance and quantification, and motivating teachers through competition (Garver 2019; Hutt and Tang 2013), value-added approaches are exemplars of neoliberal logic.

CONTEXTUAL VARIATION IN TEACHER CONTROL

Aimed at estimating individual teachers' influence on students' achievement gains, value-added approaches use multivariate regression modeling to control for factors that influence

learning but are outside teachers' control (Podgursky 2004). Dreeben (1970) described instruction, motivation, and classroom control as primary areas of teaching expertise. With teachers trained in specific strategies to improve instruction, value-added scores *should* vary based on educational attainment, degree major, teaching certification, and years of experience. On the other hand, given that they are defined as reliable stable estimates of teacher quality, value-added scores should not vary based on differences across students and schools. As a first step in accounting for factors that influence student achievement but are outside teachers' control, all value-added approaches control for students' baseline achievement (Harris 2011). Consistent with the long-standing sociological evidence on the centrality of student background for the experience of teaching and achievement (Dreeben 1970; Reardon 2011), most value-added methodologies also include controls for student social background (Milanowski 2011). EVAAS counters such controls are unnecessary because achievement *gains* should not be influenced by student background (Sanders et al. 2009b).

"Classroom viability" describes contextual variation in factors that facilitate or inhibit teachers' ability to motivate students and increase their knowledge (Dreeben 2003). Classroom viability is first affected by within-school stratification of youth, with students sorted through ability grouping and within-subject sequencing (Carbonaro 2005). In addition to differences in previous learning, lower achieving students exhibit attitudes and behaviors that impede current learning (Jennings and DiPrete 2010), particularly when clustered with other low-achieving students (Carbonaro 2005). Lower-achieving contexts include a disproportionate share of racial minority youth and youth with low socioeconomic status (SES) (Carbonaro 2005). With SES a central predictor of achievement (Reardon and Portilla 2016) and 11% of White, 27% of

Hispanic,[1] and 31% of Black children living in poverty in 2016 (Wilson and Schieder 2018), achievement disparities by SES and race are already established among kindergarteners (Quinn 2015). Within-school stratification may reflect racism and classism, as well as schools' attempts to meet diversely-prepared students' unique needs (Kelly 2009).

Classroom viability is also affected by cross-school stratification of youth. Just as the Coleman Report showed student achievement predicates more on the average characteristics of the school's students than on other school characteristics (like school funding or average teacher characteristics) (Coleman et al. 1966), classroom viability is fundamentally shaped by neighborhood and school segregation (Orfield et al. 2014). High-poverty neighborhoods are distressed by high crime rates, disrupted families, and social isolation from dominant culture (Orfield 2014). With heightened levels of academic disinvestment and externalizing behaviors among students (Farkas, Lleras and Maczuga 2002), teachers in high-poverty schools report more difficulties in raising achievement and weaker teacher-parent relationships (Allensworth, Ponisciak and Mazzeo 2009; Orfield et al. 2014). Classroom viability may also be a product of school resources (Johnson, Kraft and Papay 2011). Smaller classrooms, i.e., higher teacher to student ratios, improve teachers' ability to facilitate learning (Whitehurst and Chingos 2011). Higher per pupil expenditures may facilitate effective teaching (Darling-Hammond 2007). Other studies find teacher performance increases when their peers are effective teachers (e.g., more experienced) (Jackson and Bruegmann 2009).

In order to tangibly link results to how schools are evaluated in practice, and efficiently probe the circular logic of how we evaluate both teachers and schools, this study first follows

[1] Latino is increasingly preferred over Hispanic in some parts of the US but preferences remain mixed (http://www.pewhispanic.org/2012/04/04/when-labels-dont-fit-hispanics-and-their-views-of-identity/; https://www.nbcnews.com/news/latino/latinx-elitist-some-push-back-word-s-growing-use-n957036). I also use Hispanic to be consistent with the terminology used in communication between the district and parents.

individual teachers over time as they switch into schools of different "performance levels." The public tends to attribute the higher average test scores of schools that serve socially advantaged students to the school being higher quality. But with value-added scores applied as stable, reliable estimates of teacher quality, individual teacher's scores should not systematically vary as a function of switching school-performance-levels. Moreover, teachers' value-added scores are based on test score gains whereas school performance levels are based on average baseline test scores. To further assess whether these scores are biased by factors outside of teachers' control, this study specifically investigates whether any changes in the scores of mobile teachers can be attributed to changes in the characteristics of teachers' schools and students.

DATA AND METHODS

This study uses data on nearly all core teachers of grades 3-8 in a large urban majority-minority district between 2007-08 and 2012-13 (n=4,459), as well as data on their students and schools. This district is large, with approximately 200,000 students in 300 schools in an average year, and 12,000 teachers when considering all grade levels, subjects, and types of teachers. Teacher data is from the district and EVAAS, and student data is from the district and the state education agency. Schools are described with magnet indicators from the district, campus staff and finance indicators from the state education agency, and school level aggregations of teacher and student data. No school level variables have missing values. Multiple imputation is ill-suited to multiyear multilevel data but repeated measures in the data facilitate exact imputations rather than statistically estimated imputations. Missingness on teacher and student level measures is addressed by: 1) triangulating measures repeated for the same case across data files within the same school year (e.g., grade level), 2) triangulating measures repeated for the same case across school years (e.g., certification), prioritizing values from the nearest school year for time-

sensitive measures, and 3) by single or mean/mode imputation for a very small number of remaining missing values. This study excludes charter school teachers (about 3%) because the district does not collect data on most. This study also excludes teachers who could not be longitudinally tracked: 1,455 teachers in the district only one year and 467 teachers too mobile to fit into one of six school-performance-level trajectories (sensitivity analyses detailed in limitations section of Discussion assess the influence of these exclusions). Online Table 1 fully details teacher/student exclusion in the final dataset and rates of missingness across variables.

Teachers' School-Performance-Level-Trajectory

I construct a measure of *school-performance-level* by averaging each student's test scores across subjects, and then averaging students' mean test scores for each school. Stanford tests are developed by Pearson and administered across the nation. This district administers Stanford math, reading, and language arts tests to third through eighth graders and Stanford science and social studies tests to fourth through eighth graders. I also use scores from the Aprenda, a Spanish version of the Stanford tests. Like Corcoran, Jennings and Beveridge (2010), I standardize all test scores within each school year, subject, test version (English or Spanish), and grade level. To streamline, I construct a dichotomous measure of school-performance-level, with schools with average test scores below the median classified as low-performing and schools above the median classified as high-performing. Focusing on teachers whose first year in the district is in a low-performing school, the first three categories of the *school-performance-level-trajectory* indicator include teachers who a) stay in low-performing schools, or who switch into high-performing schools in b) 2008-09, 2009-10, or 2010-11, or c) 2011-12 or 2012-13. The remaining three categories follow a similar pattern but focus on teachers whose first year in the district is in a high-performing school.

Other Measures Describing Teachers

EVAAS estimates an annual value-added score for every core subject taught by each grade 3-8 teacher. EVAAS produces different versions of subject-specific value-added scores depending on teachers' performance pay category, school level (elementary or middle), grade level(s), and total number of subjects taught. Like Reardon and Robinson (2012), I first standardize scores (within subject, performance pay category, school level, grade level) to facilitate averaging each teacher's scores across subjects and comparing scores across schools. Score meaning was still inconsistent, though. The district financially rewards teachers with a score above the median, but EVAAS specified a different value marking the median for each score version. By comparing the value EVAAS indicated as the median on the original unstandardized variable to equivalent values on the standardized variables, I re-center all standardized scores so that zero consistently marks the median.

Decomposition-mediation analyses predict *changes in teachers' scores* with an indicator measuring the difference in each teacher's average value-added score between their last and first years in the district between 2007-08 and 2012-13. Teachers' *years of experience* is measured with a categorical variable. Dichotomous variables indicate whether teachers have a *Master's degree or higher*, and whether *teachers' college major(s)* relate to math, reading or language arts, science, social studies, education, or 'other.' This study measures *teacher certification focus* rather than certification status because all core teachers in this district are certified, and the district does not consistently document alternative certifications. Measures of teacher *race* and *gender* are included because previous studies link these qualities to teacher effectiveness for some students (Dee 2005). This study uses statistical techniques that account for teachers' time invariant qualities. Teacher qualities that vary over time and motivate mobility—namely, value-

added scores and experience—are potential confounders. *Last year's score relative to the school mean score* is the difference between each teacher's average value-added score and the mean score across teachers at their school. *Last year's experience relative to district average* subtracts each teacher's years of experience from the district mean.

Other Measures Describing Schools

Schools are first described through aggregations of student level measures described below (e.g., *proportion White*). Schools are also differentiated depending on whether they have a *magnet program*, their *per-pupil expenditure*, *teacher to student ratio* across the entire school and averaged across classrooms, average *years of experience* for all teachers and all core teachers, average *salaries* of each school's teachers with one to five years, and twenty or more years experience. For multivariate analyses, all continuous school-level predictors are standardized within each school year for increased comparability.

Measures Describing Students

Parents report their child's *gender* and *race* to the district. The district classifies each student's *economic status* with four categories: 1) not economically disadvantaged, 2) eligible for the reduced lunch program, 3) eligible for the free lunch program, or 4) living in poverty. District records indicate whether each student is in *special education*, the *Gifted and Talented program*, or an *English Learner program*. Test score data documents student *grade level* and whether the test was in *Spanish*. I construct teacher-level measures describing teachers' students because the dependent variable (teachers' value-added scores) must be the lowest level of analysis in multilevel models. Five to twenty percent of students are taught the same subject by multiple teachers (details in Online Table 1). To facilitate EVAAS' estimation of value-added scores, the district required teachers to report the degree to which they were responsible for each

student's instruction in a subject. To accurately characterize teachers' students, I link a teacher responsible for 50% of a student's instruction in math to the student's test score (for instance) times 50%. I aggregate subject-specific student-level measures to teacher-level proportions and then average across subjects. I standardize cross-subject teacher-level means within each school year.

Analytic Plan

Changes in the Value-Added Scores of Mobile Teachers. Descriptive statistics show differences in the characteristics of the teachers and students in the district's lowest and highest performing schools in 2007-08 and 2012-13. Descriptives are only shown for the first and last year of data in this study but patterns are consistent across all years. To first examine whether teachers' value-added scores are contextually variable, a growth model with the intercept fixed at the teacher level predicts teacher's average value-added score for each school year with just the school-performance-level measure. I show results from this model graphically with predicted means for each school year and for each category of teacher along the school-performance-level-trajectory measure.

Contextual Explanations for Changes in the Scores of Mobile Teachers. To determine whether any estimated effect of switching school-performance-levels on teachers' scores is mediated by changes in students and schools, I use a decomposition-mediation method based in regression modeling that was developed by Kohler, Karlson and Holm (2011). This technique produces percentage estimates of each mediator's contribution to the effect of teacher school mobility on score change, controlling on other mediators and controls. Percentage estimates express effect size and enable comparisons better than regression coefficients. Models, focused on teachers' first and last years in the district between 2007-08 and 2012-13, predict the change

in teachers' value-added scores with a dichotomous indicator of whether they switched school-performance levels. A first model focuses on teachers who switched rather than stayed in a low-performing school, and a second model on teachers who switched rather than stayed in a high-performing school. Exploratory analyses showed the results of switching are uniform regardless of the year of the switch. Measures of changes in teachers' students and schools (the potential mediators) reflect changes between teachers' first and last years in the district and are constructed with standardized variables for comparability across time. Too stable to consider as mediators, the model includes teacher characteristics as controls. To facilitate interpretation of mediation results, this table also includes descriptive statistics on each potential mediator's baseline relationship with teachers' score changes, and average differences in each potential mediator between teachers who stayed and switched.

Selection Bias Sensitivity Analyses. The models used in main analyses, fixed-intercept models, control for all time-invariant differences across teachers. But teacher qualities that vary over time and relate to mobility are potential confounders. For instance, it is possible individual teachers' value-added scores change over time as a result of selection bias rather than scores biased by contextual factors. Teachers are more likely to leave schools in which they are higher quality than their peer teachers, gravitating to schools with other more effective teachers (Feng and Sass 2017). Conversely, teachers with lower value-added scores may be pushed into lower-performing schools or courses (Glazerman et al. 2011). Inexperience is a relatively reliable measure of lower teacher efficacy (Harris and Sass 2011), and predicts teacher mobility across schools (Allensworth, Ponisciak and Mazzeo 2009). Low-performing schools typically have higher proportions of inexperienced teachers (Clotfelter, Ladd and Vigdor 2011), just as inexperienced teachers often receive the most challenging assignments within schools (Feng

2010). Whereas inexperienced teachers may seek a less challenging work environment (Allensworth, Ponisciak and Mazzeo 2009), experienced teachers may be encouraged by administrators to switch into low-performing schools (Berry 2004). In all, it is important to assess whether changes in mobile teachers' value-added scores are actually attributable to these teachers' initial value-added scores and years of experience. A first set of sensitivity analyses uses multilevel models, with the intercept fixed at the teacher level, to assess whether any relationship between context and teachers' value-added scores persists with controls for teachers' relative-score and relative-experience from the previous school year (Online Table 4). A second set of sensitivity analyses uses propensity score techniques to examine whether the changes in teachers' value-added scores and their final value-added scores are actually attributable to differences in the initial value-added scores and experience of mobile teachers (Online Tables 6-10). Propensity score analyses explained in more detail in Online Appendix A.

RESULTS

Differences by School-Performance-Level

Insert Table 1 About Here

Tables 1 through 3 provide descriptive statistics on differences by school-performance-level in the district's teachers, school resources, and students in 2007-08 and 2012-13 (the tail ends of this study's timespan). The third and sixth columns of numbers quantify the differences by school-performance-level, with a percent-change used for continuous measures and a percentage-point-change for categorical measures. Bolded values represent relatively substantively larger differences (greater than 150% or 10 percentage points). In Table 1, the value-added scores of teachers in high-performing schools are from 160% to 189% (or around 0.30 to 0.40 standard deviations (SDs)) higher than those of teachers in low-performing schools.

This is consistent with previous studies (Rice 2010; Sass et al. 2012). Teachers in low-performing schools are actually more likely than teachers in high-performing schools to have Master's degrees. The proportion of teachers with 3 years of experience or less is around 8 percentage points lower in high- relative to low-performing schools. Although teacher race is not considered an indicator of effectiveness, the proportion of teachers who are White is nearly 20 percentage points higher in high- than in low-performing schools, just as the proportion of teachers who are Black is nearly 20 percentage points higher in low- relative to high-performing schools. In all, the individual teacher qualities emphasized in neoliberal perspectives—education, certification, experience—do not vary substantially by school-performance-level.

Insert Table 2 About Here

Table 2 focuses on differences by school-performance-level in school resources. High-performing schools (around 50%) are much more likely to have a magnet program than low-performing schools (around 30%). Teachers in high-performing schools have more average years of experience than teachers in low-performing schools. The teacher to student ratio across classrooms is actually around 25% lower in high- relative to low-performing schools. Table 3 shows differences by school-performance-level in the characteristics of students. By design of the measure of school-performance-level measure, students' average test scores are much higher (around 220% higher) in high- than low-performing schools. There are also substantively large differences by school-performance-level in student economic status and race. High-performing schools serve much lower proportions of poor students – and much higher proportions of not economically disadvantaged students (differences of about 20 percentage points). The proportion of students enrolled in GT is around 15 percentage points higher in high- than in low-performing

schools. Students in high-performing schools are much more likely to be White and much less likely to be Black than students in low-performing schools.

Insert Table 3 About Here

Figure 1 shows changes over time in teachers' predicted mean value-added scores depending on their school-performance-level-trajectory. Across each school year, the average scores of teachers who consistently work in low-performing schools are around 0.25 standard deviations (SDs) lower than the district average, whereas the scores of teachers who consistently work in high-performing schools are around 0.25 SDs higher than the district average. In contrast, the average scores of teachers who switch into high-performing schools increase, just as the scores of teachers who switch into low-performing schools decrease, with the changes generally coinciding with the year of the switch. These results may indicate EVAAS scores partially reflect factors outside of teacher control.

Insert Figure 1 About Here

Contextual Factors that Mediate Score Changes of Mobile Teachers

Table 4 shows results from decomposition-mediation analyses to determine if the effect of switching school-performance-levels on changes in teachers' scores is actually mediated by contextual changes, that is, by changes in the characteristics of teachers' schools and students. As a first step in interpreting mediation, it is important to understand baseline relationships between the treatment, potential mediators, and outcome. The first column uses correlations to establish how changes in students and schools (potential mediators) more generally relate to changes in value-added scores (outcome). For instance, the -0.06 in the first cell of this column indicates teachers whose number of students increased experienced an average decrease in their value-added scores over time. Overall, student characteristics (at teacher- and school-level) relate

to score changes in the same way they relate to baseline achievement levels, with increases in socially advantaged students relating to increasing value-added scores over time. Similarly, increases in most school resources relate to increasing value-added scores.

Insert Table 4 About Here

To explore how switching school performance levels (the treatment) relates to potential mediators (changes in teachers' students and schools), Table 4 provides four columns showing differences-in-means. These differences-in-means show how the characteristics of teachers' students and schools by the end of the study period had changed relative to the beginning of the study period. The first column of differences-in-means shows contextual changes for teachers who stayed in low-performing schools, the second column for teachers who switched from low- to high-performing, the third for teachers who stayed in high-performing schools, and the fourth for teachers who switched from high- to low-performing. Focusing on the largest contextual changes (bolded), teachers who switch into high-performing schools experience, on average, a 44-percentage-point increase in the percent of students at the school who are GT, a 41-percentage-point decrease in the percent of students in their class(es) who live in poverty, and a 38-percentage-point increase in the percent of students in their class(es) who are not economically disadvantaged. In contrast, teachers who switch into low-performing schools experience, on average, a 46-percentage-point increase in the percent of students at the school living in poverty and a corresponding 45-percentage-point decrease in the percent of students not economically disadvantaged. Similar shifts in the economic status of the students in their class(es) represent the other largest changes experienced by teachers who switch into low-performing schools.

The two columns in Table 4 with indirect effects are explicit estimates of mediation, representing the percent contribution of each contextual change to the total effect of switching school-performance-levels on changes in teachers' scores. Bolded values indicate the largest measured mediators. Hyphens indicate the characteristic did not function as a mediator (these measures still serve as controls for less biased estimates for other mediators). In the largest contributor, 7.0% of the positive effect of switching into a high-performing school is attributable to average decreases in the proportion of students at the school living in poverty. In the other largest contributors, a combined 10.9% of the positive effect of switching is explained by shifts in the characteristics of the students in these teachers' new classrooms: an increased proportion White (3.2%), an increased proportion in grade 6 (3.1%), an increased proportion not economically disadvantaged (2.5%), and a decreased proportion living in poverty (2.1%). For teachers who switch into low-performing schools, a combined 8.2% of the negative effect of switching on their scores is attributable to shifts in the characteristics of the students in their classrooms: increased proportion living in poverty (2.2%), increased proportion Black (1.7%), decreased proportion eighth graders (1.6%), increased proportion English Learners (1.4%), and decreased proportion in GT (1.3%). These results suggest EVAAS scores partially capture contextual influences beyond the control of teachers.

DISCUSSION

Value-added scores build on neoliberal ideals by framing teachers' work and motivations individualistically. This framing runs counter to evidence on the importance of contextual factors for teachers' work and students' learning. Fixed-intercept models tracking individual teachers between 2007-08 and 2012-13 show scores increase for teachers who switched into high-performing schools, and decrease for teachers who switch into low-performing schools.

Particularly indicative of scores biased by contextual factors outside teachers' control, results from mediation analyses show score changes for mobile teachers are partially attributable to shifts in the economic status and race of students in teachers' classrooms and schools. Goldhaber and Hansen (2013) also found temporal inconsistencies in teachers' scores but did not identify causes. Jackson (2013) found school mobility benefitted scores but did not differentiate by school-performance-level. The following paragraphs expand on the implications of neoliberalism for teachers and children, as well as the more broad implications of neoliberalism for societal problems and the laborers taxed with resolving those problems.

The Duality of Neoliberalism

Consistent with neoliberal duality, EVAAS defends their approach as a refusal to have lower expectations for socially disadvantaged children (Sanders et al. 2009b), which is essentially an admission that their method does not account for how inequality outside of schools affects the work of teachers. The centrality of structural factors like poverty for education outcomes is well-documented, with differences across schools and teachers explaining less than a quarter of the variation in achievement (Gamoran and Long 2006; Hill 2016). Accountable actors are, by definition, autonomous and self-determining, yet curricular standards and classroom approaches are increasingly standardized, leaving teachers with less control over their classrooms. Renzulli (2014:149) described this as "simultaneous and contradictory transformations in decentralization and standardization." Neoliberalism is often counterintuitively marked by an expansion of state (Pacewicz 2013), sometimes specifically aimed at ensuring public institutions fail (Valocchi 2017) [e.g., the Affordable Care Act (Redhead and Kinzer 2013)]. Although professionals within public institutions are generally not

well-positioned for success (Gillborn and Youdell 1999), state legitimacy depends on public sector failure appearing natural.

Naturalizing Neoliberal Perspectives

The propagation of neoliberalism depends on the naturalization of its fundamental concepts (Harvey 2006). Explanations for undesirable outcomes that run counter to neoliberal perspectives, such as social inequality, are systematically discounted and suppressed (Trumpy 2008). Intersectionalists document how neoliberalism perpetuates beliefs that inequalities are sourced in individual failings (Hill Collins and Bilge 2016). Similarly, some researchers use lower average value-added scores to argue low-performing schools are the result of low quality teachers (Rice 2010; Sass et al. 2012), which is circular reasoning given this study's findings. Consistent with other studies (Hanushek and Rivkin 2012), supplementary analyses showed the scores of teachers in this district vary more within- than between-schools, suggesting differences in teaching are not a good explanation for differences in school performance levels (Croninger et al. 2007). Transparent and specific descriptions of EVAAS methodology are unavailable [e.g., Wright et al. (2010)], with the aim, some argue, of protecting the for-profit methodology from duplication and criticism (Corcoran 2010). EVAAS argues their use of linear mixed models, "algorithms [which] have been employed for decades," negate the need for peer review (Sanders et al. 2009a). Neoliberal approaches use quantification to promote an illusion of meritocratic evaluation (Mehta and Davies 2018), to reify teachers and schools as the source of disadvantaged students' lower test score gains.

Although this study's results cannot be generalized to other value-added approaches or beyond this district, only a minority of previous studies claim unbiased scores are possible (Bacher-Hicks, Kane and Staiger 2014; Bau and Das 2016; Chetty, Friedman and Rockoff 2014).

A much larger number of previous studies—none focused on EVAAS—criticize value-added approaches for insufficiently accounting for the nonrandom sorting of teachers and students across classrooms (Goldhaber, Gabele and Walch 2012; Hill, Kapitula and Umland 2011; McCaffrey et al. 2004; McCaffrey et al. 2003; Newton et al. 2010; Rothstein 2010), and for being unreliable across different models (Sass, Semykina and Harris 2014), tests (Ballou and Springer 2015; Lockwood et al. 2006; Papay 2011; Polikoff 2014), years (Darling-Hammond et al. 2011; Mariano, McCaffrey and Lockwood 2010; McCaffrey et al. 2009), and subjects (Goldhaber, Cowan and Walch 2013). Although various school districts face lawsuits related to charges of biased scores (Pallas 2016; Strauss 2014), and a teacher's suicide in Los Angeles was linked to the public release of value-added scores (National Education Policy Center 2018), the use of value-added approaches remains prevalent.

Neoliberal policies, by design, increase occupational precariousness (Garver 2019; Kristal 2013). Teachers are already low status in the US, with salaries lower relative to comparable workers in the US (Allegretto and Mishel 2016), and relative to teachers in similarly resourced countries (OECD 2016). Neoliberal perspectives support assertions by policymakers that teachers are the cause of intra- and cross-national educational disparities (Clotfelter, Ladd and Vigdor 2010). With public university faculty increasingly facing similar scrutiny, professors' professional status is degraded by the outsourcing of instruction to online modules (Tabb 2001) and to underpaid contingent faculty (Edmonds 2015). The extent to which neoliberal perspectives are naturalized is evident in trial and appellate court narratives that emphasize teacher agency at the exclusion of the broader social context of schooling (Gottlieb, Hutt and Superfine 2018). Neoliberal approaches may not only fail to accomplish positive change but may actually harm the teaching profession.

Obstruction of Knowledge and Sound Policy

Neoliberal perspectives and approaches mask lines of knowledge essential for effective educational policy, or, for that matter, any policy aimed at the common good (Stecher et al. 2018). The policy habits of thought that attribute low-performing schools to low quality teaching eclipse the central contributions of segregation (Quiroz, Milam-Brooks and Adams-Romena 2014). Efforts toward school desegregation have stalled and reversed since the 1990s (Orfield et al. 2014). The notion that separate can be equal is propagated by portrayals of schools and teachers that are exceptions as models (Berliner and Biddle 1995). High-performing high-poverty schools are rare, and usually have some mechanism for selective enrollment (Logan, Minca and Adar 2012). Conflating the contextual factors that influence learning with teacher quality limits understanding of the nature of teaching (Ball 1994). Teaching is substantially more challenging in high-poverty schools (Ingersoll 2005), a fact eclipsed by the use of statistical models with basic administrative data to 'account' for these differences. Socioeconomic status is a key predictor of achievement across a diversity of contexts (Reardon and Portilla 2016). Although achievement gaps by class and race are evident at kindergarten and remain fairly constant across grade levels (Cheadle 2008), education policy continues to emphasize the role of schools and teachers in education disparities (Loveless 2003). These perversions of knowledge shift the burden of poverty from the state to the shoulders of teachers and the community itself (Apple 2006). Education is not the only realm increasingly marked by data collection and 'research-based evidence.' Although privileged within the neoliberal preference for quantification and precision, quantitative data and statistics are subject to bias through the theories that guide measurement, focus, and analytic assumptions (Kearns and Roth 2019).

Neoliberalism explicitly obscures knowledge and understanding, promoting an ideologically based prioritization of certain perspectives.

Limitations and Conclusion

Limitations of this study merit mention. Like all studies on contextually variable education policies, results are not generalizable to other value-added approaches or beyond this district. Nonetheless, the relative prevalence of the EVAAS approach increases the policy relevance of the study's findings. The EVAAS approach may result in more biased scores than other approaches that include more controls for student differences, but studies that claim value-added scores can be unbiased are in the minority. By taking a creative approach to investigate a widely implemented but under-studied value-added methodological approach, this study contributes to sociological understandings of teacher effects and general processes of neoliberal policies.

This study's dichotomous measure of school performance level facilitates streamlined and visually powerful findings but may obscure differences between middle-performing schools and schools at the extreme ends of performance. Sensitivity analyses show results are robust to a more nuanced measure of school performance level, with the pattern of increased value-added scores consistent across school performance level quartiles (Online Table 2). More specifically, the scores of teachers who switch from Quartile 1 schools (lowest performing) into Quartile 4 schools (highest performing) increase more than the scores of teachers who switch from Quartile 1 schools into Quartile 2 or 3 schools. Main results are also supported by sensitivity analyses that use this quartile measure of school performance level to find that increases and decreases in the value-added scores of mobile teachers are partially attributable to changes in the characteristics of their students and schools (Online Table 2).

By longitudinally tracking individual teachers, this study powerfully accounts for the potential influence of unmeasured differences across teachers. To facilitate this longitudinal tracking of individual teachers, main analyses exclude teachers with mixed-performance-level trajectories and teachers in the district one year. Supplemental descriptive statistics show that teachers included in main analyses have lower value-added scores on average than mixed-trajectory teachers and higher value-added scores than teachers in the district for just one year; patterns of advantage and disadvantage are similar along measures of these groups of teachers' educational history and experience (Online Table 3). Confidence in main analyses is increased by similar findings from sensitivity analyses that include all teachers. First, regression models with the intercept fixed at the teacher level show contextual factors relate significantly to teachers' value-added scores each year even with mixed-trajectory and single-year teachers included (Online Table 4). Second, decomposition-mediation analyses that include all teachers show contextual factors, like the proportion of students who are economically disadvantaged, mediate the estimated effect of teaching in a low- rather than high-performing school, in each school year (Online Table 5).

Selection bias is a substantial concern. Although this study's main analyses rely on fixed-intercept modeling, which controls for all unmeasured time invariant differences across teachers, the possibility remains that mobile teachers differ in other ways that influence their value-added scores. First, fixed-intercept models show that contextual factors relate to teachers' value-added scores each school year even when controlling for time-varying teacher qualities like teachers' relative score and relative experience from each previous year (Online Table 4). Second, the average increase in the scores of teachers who switch into high-performing schools, and the average decrease in the scores of teachers who switch into low-performing schools, are

robust to propensity score techniques (Online Tables 9-10). Moreover, propensity score techniques show differences in the final scores of teachers who do and do not switch school-performance-level are not explained by differences in teachers' initial experience or quality (as measured by teachers' initial value-added score) (Online Tables 9-10).

Some describe contextual variation in teachers' scores as a valid indicator of shifts in teacher efficacy (Goldhaber and Hansen 2013). Teacher effectiveness is undoubtedly contextually variable, for the reasons outlined in this study. In this case, value-added scores should be described as measures of contextual factors that facilitate effective teaching rather than measures of individual teacher quality. EVAAS's defense of their approach as a refusal to have lower expectations for socially disadvantaged children (Sanders et al. 2009b) is an ideological (and idealistic) rather than empirical approach to value-added modeling, based more in what we hope teachers could accomplish than in the realities of the structure of our society and schools. Ultimately, educational disparities will be most effectively addressed by targeting inequality in our society (Rothstein 2004). Achievement disparities are greater in countries with more class disparities (Montt 2011), and social inequality is higher in the US than similarly developed countries (Carnoy and Rothstein 2013). Although the conflation of school and student background effects with teacher effects on achievement is a common point of study for sociologists of education (Morgan and Shackelford 2018), applied instances of neoliberal logic, like value-added scores, are largely interrogated by economists rather than sociologists (Harris and Herrington 2015). This parallels the relative dominance of economists in education policy (Dworkin 2005; Schneider, Grogan and Maier 2011), and policy more broadly (Smith 2016; Wolfers 2015). Renzulli (2014) depicts this disconnect as detrimental for both policy and sociological research. Encroaching neoliberalism may even reflect and reproduce these

processes. Just as sociologists tend to attack the purposes and practices of education (Schneider 2003), Harvey (2006) describes an "ideological assault … upon educational institutions" as one of three key elements in the US's transition towards neoliberalism. Paying teachers who work in high-poverty schools a higher salary should be standard practice. Schools should be sufficiently funded to provide social services for all the issues they are expected to resolve.

REFERENCES

Allegretto, Sylvia, and Lawrence Mishel. 2016. "The Teacher Pay Gap is Wider Than Ever." Washington, DC: Economic Policy Institute.

Allensworth, Elaine, Stephen Ponisciak, and Christopher Mazzeo. 2009. "The Schools Teachers Leave: Teacher Mobility in Chicago Public Schools." Chicago, IL: Consortium on Chicago School Research, Urban Education Institute, University of Chicago.

Ambrosio, John. 2013. "Changing the Subject: Neoliberalism and Accountability in Public Education." *Educational Studies* 49(4):316-33.

Amrein-Beardsley, Audrey, and Clarin Collins. 2012. "The SAS Education Value-Added Assessment System (SAS® EVAAS®) in the Houston Independent School District (HISD): Intended and Unintended Consequences." *Education Policy Analysis Archives* 20(12):1-31.

Apple, Michael. 2006. "Understanding and Interrupting Neoliberalism and Neoconservatism in Education." *Pedagogies: An International Journal* 1(1):21-26.

Bacher-Hicks, Andrew, Thomas J. Kane, and Douglas O. Staiger. 2014. "Validating Teacher Effect Estimates Using Changes in Teacher Assignments in Los Angeles (Working Paper 20657)." Cambridge, MA: National Bureau of Economic Research.

Ball, Stephen J. 1994. *Education Reform: A Critical and Post-Structural Approach*. Berkshire, UK: Open University Press.

—. 2016. "Neoliberal Education? Confronting the Slouching Beast." *Policy Futures in Education* 14(8):1046-59.

Ballou, Dale, and Matthew G. Springer. 2015. "Using Student Test Scores to Measure Teacher Performance: Some Problems in the Design and Implementation of Evaluation Systems." *Educational Researcher* 44(2):77-86.

Bau, Natalie, and Jishnu Das. 2016. "The Misallocation of Pay and Productivity in the Public Sector: Evidence from the Labor Market for Teachers (Working Paper)."

Berliner, David C., and Bruce C. Biddle. 1995. *The Manufactured Crisis: Myths, Fraud, and the Attack on America's Public Schools*. Reading, MA: Addison-Wesley Publishing Company, Inc.

Berry, Barnett. 2004. "Recruiting and Retaining "Highly Qualified Teachers" for Hard-to-Staff Schools." *NASSP Bulletin* 88(638):5-27.

Bidwell, Charles. 1965. "The School as a Formal Organization." in *Handbook of Organizations*, edited by James March. Chicago: Rand McNally.

Carbonaro, William. 2005. "Tracking, Students' Effort, and Academic Achievement." *Sociology of Education* 78(1):27-49.

Carnoy, Martin, and Richard Rothstein. 2013. *What Do International Tests Really Show About U.S. Student Performance?* Palo Alto, CA: Economic Policy Institute.

Cheadle, Jacob E. 2008. "Educational Investment, Family Context, and Children's Math and Reading Growth from Kindergarten Through the Third Grade." *Sociology of Education* 81(1):1-31.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review* 104(9):2593-632.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2010. "Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects." *Journal of Human Resources* 45(3):655-81.

—. 2011. "Teacher Mobility, School Segregation, and Pay-Based Policies to Level the Playing Field (Working Paper 44)." *Education Finance and Policy* 6(3):399-438.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederick D. Weinfall, and Robert L. York. 1966. *Equality of Educational Opportunity*. Washington, D.C.: National Center for Educational Statistics.

Corcoran, Sean P. 2010. "Can Teachers Be Evaluated by their Students' Test Scores? Should They Be? The Use of Value-Added Measures of Teacher Effectiveness in Policy and Practice." in *Education Policy for Action Series*. Providence, Rhode Island: Annenberg Institute for School Reform, Brown University.

Corcoran, Sean P., Jennifer L. Jennings, and Andrew A. Beveridge. 2010. "Teacher Effectiveness on High- and Low-Stakes Tests (Working Paper)." New York, NY: New York University.

Croninger, Robert G., Jennifer King Rice, Amy Rathbun, and Masako Nishio. 2007. "Teacher Qualifications and Early Learning: Effects of Certification, Degree, and Experience on First-Grade Student Achievement." *Economics of Education Review* 26:312-24.

Darling-Hammond, Linda. 2007. "Race, Inequality and Educational Accountability: The Irony of 'No Child Left Behind'." *Race Ethnicity and Education* 10(3):245-60.

Darling-Hammond, Linda, Audrey Amrein-Beardsley, Edward H. Haertel, and Jesse Rothstein. 2011. *Getting Teacher Evaluation Right: A Background Paper for Policy Makers (A Research Briefing)*. Washington, DC: American Educational Research Association, National Academy of Education.

Dee, Thomas S. 2005. "A Teacher Like Me: Does Race, Ethnicity, or Gender Matter?" *American Economic Review* 95(2):158-65.

Dreeben, Robert. 1970. *The Nature of Teaching: Schools and the Work of Teachers*. Glenview, IL: Scott Foresman.

—. 2003. "Classrooms and Politics." Pp. 229-25- in *Stability and Change in American Education: Structure, Process, and Outcomes*, edited by Maureen T. Hallinan, Adam Gamoran, Warren Kubitschek, and Tom Loveless. Clinton Corners, NY: Eliot Werner Publications, Inc.

Dworkin, Gary A. 2005. "The No Child Left Behind Act: Accountability, High-Stakes Testing, and Roles for Sociologists." *Sociology of Education* 78(2):170-74.

Edmonds, Dan. 2015. "More Than Half of College Faculty are Adjuncts: Should You Care?" in *Forbes*. Jersey City, NJ: Forbes Media LLC.

Farkas, George, Christy Lleras, and Steve Maczuga. 2002. "Does Oppositional Culture Exist in Minority and Poverty Peer Groups?" *American Sociological Review* 67(1):148-55.

Feng, Li. 2010. "Hire Today, Gone Tomorrow: New Teacher Classroom Assignments and Teacher Mobility." *Education Finance and Policy* 5(3):278-316.

Feng, Li, and Tim Sass. 2017. "Teacher Quality and Teacher Mobility." *Education Finance and Policy* 12(3):396-418.

Finster, Matthew, and Jackson Miller. 2014. *Who's at the Top? It Depends... Identification of Highest Performing Schools for Newsweek's 2014 High School Rankings.* Rockville, MD: Westat.

Gamoran, Adam, and Daniel A. Long. 2006. *Equality of Educational Opportunity: A 40-Year Retrospective (WCER Working Paper No. 2006-9).* Madison, WI: University of Wisconsin-Madison.

Garver, Rachel. 2019. "Evaluative Relationships: Teacher Accountability and Professional Culture." *Journal of Education Policy* Published Online First.

Gillborn, David, and Deborah Youdell. 1999. *Rationing Education: Policy, Practice, Reform and Equity*. Berkshire, UK: Open University Press.

Glazerman, Steven, Hanley Chiang, Alison J. Wellington, Jill M. Constantine, and Daniel Player. 2011. *Impacts of Performance Pay Under the Teacher Incentive Fund: Study Design Report*. Washington, DC: Mathematica Policy Research.

Goldhaber, Dan, James Cowan, and Joe Walch. 2013. "Is a Good Elementary Teacher Always Good? Assessing Teacher Performance Estimates Across Subjects." *Economics of Education Review* 36:216-28.

Goldhaber, Dan, Brian Gabele, and Joe Walch. 2012. *Does the Model Matter? Exploring the Relationship Between Different Achievement-Based Teacher Assessments (CEDR Working Paper 2012-6).* Seattle, WA: University of Washington.

Goldhaber, Dan, and Michael Hansen. 2013. "Is it Just a Bad Class? Assessing the Long-term Stability of Estimated Teacher Performance." *Economica* 80(319):589-612.

Goodman, Philip. 2012. "“Another Second Chance” Rethinking Rehabilitation through the Lens of California's Prison Fire Camps." *Social Problems* 59(4):437-58.

Gottlieb, Jessica J., Ethan L. Hutt, and Benjamin M. Superfine. 2018. "Causal Stories in Vergara v. California." *Educational Policy* Published online first.

Hanushek, Eric A., and Steven G. Rivkin. 2012. "The Distribution of Teacher Quality and Implications for Policy." *Annual Review of Economics* 4:131-57.

Harris, Douglas N. 2011. "Value-Added Measures and the Future of Educational Accountability." *Science* 333(6044):826-27.

Harris, Douglas N., and Carolyn D. Herrington. 2015. "Editors' Introduction: The Use of Teacher Value-Added Measures in Schools: New Evidence, Unanswered Questions, and Future Prospects." *Educational Researcher* 44(2):71-76.

Harris, Douglas N., and Tim R. Sass. 2011. "Teacher Training, Teacher Quality, and Student Achievement." *Journal of Public Economics* 95:798-812.

Harvey, David. 2006. "Neo-Liberalism as Creative Destruction." *Geografiska Annaler. Series B, Human Geography* 88(2):145-58.

Hill Collins, Patricia, and Sirma Bilge. 2016. *Intersectionality*. Cambridge, UK: Polity.

Hill, Heather C. 2016. "50 Years Ago, One Report Introduced Americans to the Black-White Achievement Gap. Here's What We've Learned Since." in *Chalkbeat*. New York, NY: Chalkbeat.

Hill, Heather C., Laura Kapitula, and Kristin Umland. 2011. "A Validity Argument Approach to Evaluating Teacher Value-Added Scores." *American Educational Research Journal* 48(3):794-831.

Hong, Guanglei, and Stephen W. Raudenbush. 2005. "Effects of Kindergarten Retention Policy on Children's Cognitive Growth in Reading and Mathematics." *Educational Evaluation and Policy Analysis* 27(3):205-24.

Hutt, Ethan, and Aaron Tang. 2013. "The New Education Malpractice Litigation." *Virginia Law Review* 99(3):419-92.

Ingersoll, Richard M. 2005. "The Anomaly of Educational Organizations and the Study of Organizational Control." Pp. 91-110 in *The Social Organization of Schooling*, edited by Larry V. Hedges and Barbara Schneider. New York, NY: Russell Sage Foundation.

Jackson, C. Kirabo. 2013. "Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence from Teachers." *The Review of Economics and Statistics* 95(4):1096-116.

Jackson, C. Kirabo, and Elias Bruegmann. 2009. "Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers." *American Economic Journal: Applied Economics* 1(4):85-108.

Jennings, Jennifer L., and Thomas A. DiPrete. 2010. "Teacher Effects on Social and Behavioral Skills in Early Elementary School." *Sociology of Education* 83(2):135-59.

Johnson, Susan Moore, Matthew A. Kraft, and John P. Papay. 2011. "How Context Matters in High-Need Schools: The Effects of Teachers' Working Conditions on Their Professional Satisfaction and Their Students' Achievement." in *ASHE Higher Education Report*. Cambridge, MA: Project on the Next Generation of Teachers, Harvard Graduate School of Education.

Kearns, Michael, and Aaron Roth. 2019. *The Ethical Algorithm: The Science of Socially Aware Algorithm Design*. Oxford, UK: Oxford University Press.

Kelly, Sean. 2009. "The Black-White Gap in Mathematics Course Taking." *Sociology of Education* 82(1):47-69.

Kohler, Ulrich, Kristian Bernt Karlson, and Anders Holm. 2011. "Comparing Coefficients of Nested Nonlinear Probability Models." *The Stata Journal* 11(3):420-38.

Kristal, Tali. 2013. "Slicing the Pie State Policy, Class Organization, Class Integration, and Labor's Share of Israeli National Income." *Social Problems* 60(1):100-27.

Kupermintz, Haggai. 2003. "Teacher Effects and Teacher Effectiveness: A Validity Investigation of the Tennessee Value Added Assessment System." *Educational Evaluation and Policy Analysis* 25(3):287-98.

Lockwood, J.R., Daniel F. McCaffrey, Laura S. Hamilton, Brian Stecher, Vi-Nhuan Le, and Felipe Martinez. 2006. *The Sensitivity of Value-Added Teacher Effect Estimates to Different Mathematics Achievement Measures*. Santa Monica, CA: The Rand Corporation.

Logan, John R., Elisabeta Minca, and Sinem Adar. 2012. "The Geography of Inequality: Why Separate Means Unequal in American Public Schools." *Sociology of Education* 85(3):287-301.

Loveless, Tom. 2003. "The Regulation of Teaching and Learning." Pp. 171-92 in *Stability and Change in American Education: Structure, Process, and Outcomes*, edited by Maureen T. Hallinan, Adam Gamoran, Warren Kubitschek, and Tom Loveless. Clinton Corners, NY: Eliot Werner Publications, Inc.

Mariano, Louis T., Daniel F. McCaffrey, and J. R. Lockwood. 2010. "A Model for Teacher Effects from Longitudinal Data Without Assuming Vertical Scaling." *Journal of Educational and Behavioral Statistics* 35(3):253-79.

McCaffrey, Daniel F., and Laura S. Hamilton. 2007. "Value-Added Assessment in Practice: Lessons from the Pennsylvania Value-Added Assessment System Pilot Project." Santa Monica, CA: The Rand Corporation.

McCaffrey, Daniel F., J. R. Lockwood, Daniel Koretz, Thomas A. Louis, and Laura Hamilton. 2004. "Models for Value-Added Modeling of Teacher Effects." *Journal of Educational and Behavioral Statistics* 29(1):67-101.

McCaffrey, Daniel F., J. R. Lockwood, Daniel M. Koretz, and Laura S. Hamilton. 2003. *Evaluating Value-Added Models for Teacher Accountability*. Santa Monica, CA: RAND Corporation.

McCaffrey, Daniel F., Tim R. Sass, J. R. Lockwood, and Kata Mihaly. 2009. "The Intertemporal Variability of Teacher Effect Estimates." *Education Finance and Policy* 4(4):572-606.

Mehta, Jal, and Scott Davies. 2018. "Education in a New Society: Renewing the Sociology of Education." Pp. 1–58 in *Education in a New Society: Renewing the Sociology of Education*, edited by Jal Mehta and Scott Davies. Chicago, IL: Chicago University Press.

Meyer, Heinz-Dieter, and Brian Rowan. 2006. "Institutional Analysis and the Study of Education." Pp. 1-13 in *The New Institutionalism*, edited by Heinz-Dieter Meyer and Brian Rowan. Albany, NY: SUNY Press.

Milanowski, Anthony. 2011. *Resolving Some Issues in Using Value-Added Measures of Productivity for School and Teacher Incentives: Ideas from Technical Assistance and TIF Grantee Experience*. Washington, DC: U.S. Department of Education, Office of Elementary and Secondary Education, Center for Educator Compensation Reform.

Montt, Guillermo. 2011. "Cross-National Differences in Educational Achievement Inequality." *Sociology of Education* 84(1):49-68.

Morgan, Stephen L., and Daniel T. Shackelford. 2018. "School and Teacher Effects." in *Handbook of the Sociology of Education in the 21st Century*, edited by Barbara Schneider. New York, NY: Springer International Publishing.

Nancarrow, Susan A., and Alan M. Borthwick. 2005. "Dynamic Professional Boundaries in the Healthcare Workforce." *Sociology of Health & Illness* 27(7):897-919.

National Education Policy Center. 2018. *Whatever Happened with the Los Angeles Times' Decision to Publish Teachers' Value-Added Scores?* Boulder, CO: National Education Policy Center, University of Colorado Boulder School of Education.

Newton, Xiaoxia A., Linda Darling-Hammond, Edward Haertel, and Ewart Thomas. 2010. "Value-Added Modeling of Teacher Effectiveness: An Exploration of Stability across Models and Contexts." *Education Policy Analysis Archives* 18(23):2-27.

OECD. 2016. *Education at a Glance 2016: OECD Indicators*. Paris, France: Organisation for Economic Co-operation Development.

Orfield, Gary. 2014. "Tenth Annual *Brown* Lecture in Education Research: A New Civil Rights Agenda for American Education." *Educational Researcher* 43(6):273-92.

Orfield, Gary, Erica Frankenberg, Jongyeon Ee, and John Kuscera. 2014. *Brown at 60: Great Progress, a Long Retreat and an Uncertain Future*. Los Angeles, CA: The Civil Rights Project, University of California, Los Angeles.

Pacewicz, Josh. 2013. "Regulatory Rescaling in Neoliberal Markets." *Social Problems* 60(4):433-56.

Pallas, Aaron. 2016. "Teacher Evaluations Fail a Crucial Test: A Court Rules that the Piece of Teacher's Rating Determined by Student Test Scores is 'Arbitrary and Capricious'." in *New York Daily News*. New York, NY: New York Daily News.

Papay, John P. 2011. "Different Tests, Different Answers: The Stability of Teacher Value-Added Estimates across Outcome Measures." *American Educational Research Journal* 48(1):163-93.

Piketty, Thomas, and Emmanuel Saez. 2004. "Income Inequality in the United States, 1913-2002." Pp. 141-225 in *Top Incomes over the Twentieth Century*, edited by A. B. Atkinson and Thomas Piketty. New York, NY: Oxford University Press.

Podgursky, Michael. 2004. *Comments on "How to Improve the Supply of High-Quality Teachers" by Eric A. Hanushek and Steven G. Rivkin*. Washington, DC: Brookings Institution Press.

Polikoff, Morgan S. 2014. "Evaluating Teachers When Tests Differ in Their Sensitivity to Instruction." Pp. 278-302 in *Designing Teacher Evaluation Systems-New Guidance from the Measures of Effective Teaching Project*, edited by Thomas J. Kane, Kerri A. Kerr, and Robert C. Pianta. San Francisco, CA: Jossey-Bass.

Quinn, David M. 2015. "Kindergarten Black–White Test Score Gaps: Re-examining the Roles of Socioeconomic Status and School Quality with New Data." *Sociology of Education* 88(2):120-39.

Quiroz, Pamela Anne, Kisha Milam-Brooks, and Dominique Adams-Romena. 2014. "School as Solution to the Problem of Urban Place: Student Migration, Perceptions of Safety, and Children's Concept of Community." *Childhood* 21(2):207-25.

Reardon, Sean F. 2011. "The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations." in *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*. New York, NY: Russell Sage Foundation.

Reardon, Sean F., and Ximena A. Portilla. 2016. *Recent Trends in Income, Racial, and Ethnic School Readiness Gaps at Kindergarten Entry (CEPA Working Paper No. 15-02)*. Stanford, CA: Center for Education Policy Analysis, Stanford University.

Reardon, Sean F., and Joseph P. Robinson. 2012. "Regression Discontinuity Designs with Multiple Rating-Score Variables." *Journal of Research on Educational Effectiveness* 5(1):83-104.

Redhead, C. Stephen, and Janet Kinzer. 2013. *Legislative Actions to Repeal, Defund, or Delay the Affordable Care Act*. Washington, DC: Congressional Research Service, Library of Congress.

Reich, Jennifer A. 2014. "Neoliberal Mothering and Vaccine Refusal: Imagined Gated Communities and the Privilege of Choice." *Gender & Society* 28(5):679-704.

Renzulli, Linda. 2014. "Educational Transformations and Why Sociology Should Care." *Social Currents* 1(2):149-56.

Rice, Jennifer King. 2010. *The Impact of Teacher Experience: Examining the Evidence and Policy Implications (Brief 11)*. Washington, DC: National Center for Analysis of Longitudinal Data in Education Research, American Institutes for Research.

Robinson-Cimpian, Joseph P. 2016. "Propensity Score Matching." in *AERA Institute on Statistical Analysis*. Washington, DC.

Rothstein, Jesse. 2010. "Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement." *The Quarterly Journal of Economics* 125(1):175-214.

Rothstein, Richard. 2004. *Class and Schools: Using Social, Economic, and Educational Reform to Close the Black-White Achievement Gap*. New York, NY: Teachers College Press.

Sanders, William L., S. Paul Wright, June C. Rivers, and Jill G. Leandro. 2009a. *Addressing Common Concerns About Value-Added Modeling (White Paper)*. Cary, NC: SAS.

—. 2009b. *A Response to Criticisms of SAS EVAAS*. Cary, NC: SAS Institute, Inc.

Sass, Tim, Jane Hannaway, Zeyu Xu, David Figlio, and Li Feng. 2012. "Value Added of Teachers in High-Poverty Schools and Lower Poverty Schools." *Journal of Urban Economics* 72(1):104-22.

Sass, Tim R., Anastasia Semykina, and Douglas N. Harris. 2014. "Value-Added Models and the Measurement of Teacher Productivity." *Economics of Education Review* 38:9-23.

Schneider, Barbara. 2003. "Sociology of Education: An Overview of the Field at the Turn of the Twenty-First Century." Pp. 193-228 in *Stability and Change in American Education: Structure, Process, and Outcomes*, edited by Maureen T. Hallinan, Adam Gamoran, Warren Kubitschek, and Tom Loveless. Clinton Corners, New York: Eliot Werner Publications, Inc.

Schneider, Barbara, Erin Grogan, and Adam Maier. 2011. "Improving Teacher Quality: A Sociological Presage." Pp. 163-80 in *Frontiers in Sociology of Education, Frontiers in Sociology and Social Research*, edited by Maureen T. Hallinan. Netherlands: Springer.

Silva, Jennifer M. 2019. *We're Still Here: Pain and Politics in the Heart of America*. Oxford, UK: Oxford University Press.

Smith, Noah. 2016. "Calling All Sociologists: America Needs You." in *Bloomberg View*. New York, NY: Bloomberg.

Stecher, Brian M., Deborah J. Holtzman, Michael S. Garet, Laura S. Hamilton, John Engberg, Elizabeth D. Steiner, Abby Robyn, Matthew D. Baird, Italo A. Gutierrez, Evan D. Peet, Iliana Brodziak De Los Reyes, Kaitlin Fronberg, Gabriel Weinberger, Gerald Paul Hunter, and Jay Chambers. 2018. *Improving Teaching Effectiveness: Final Report - The Intensive Partnerships for Effective Teaching Through 2015-2016.* Santa Monica, CA: Rand Corporation.

Strauss, Valerie. 2014. "High-Achieving Teacher Sues State Over Evaluation Labeling Her "Ineffective"." in *Washington Post*. Washington DC: The Washington Post.

Tabb, William K. 2001. "Essay: Globalization and Education as a Commodity." in *PSC Clarion*. New York, NY: Professional Staff Congress, City University of New York.

Trumpy, Alexa J. 2008. "Subject to Negotiation: The Mechanisms Behind Co-Optation and Corporate Reform." *Social Problems* 55(4):480-500.

Valocchi, Stephen. 2017. "Capitalisms and Gay Identities: Towards a Capitalist Theory of Social Movements." *Social Problems* 64(2):315-31.

Vasseur, Michael. 2016. "Incentives or Mandates? Determinants of the Renewable Energy Policies of U.S. States, 1970-2012." *Social Problems* 63(2):284-301.

Vosters, Kelly N., Cassandra M. Guarino, and Jeffrey M. Wooldridge. 2018. *Understanding and Evaluating the SAS EVAAS Univariate Response Model (URM) for Measuring Teacher Effectiveness (Working Paper No. 2018-001).* Charlotte, NC: The University of North Carolina at Charlotte.

Weick, Karl E. 1976. "Educational Organizations as Loosely Coupled Systems." *Administrative Science Quarterly* 21(1):1-19.

Whitehurst, Grover J., and Matthew M. Chingos. 2011. *Class Size: What Research Says and What it Means for State Policy*. Washington, DC: Brown Center on Education Policy at Brookings.

Wilson, Valerie, and Jessica Schieder. 2018. *The Rise in Child Poverty Reveals Racial Inequality, More than a Failed War on Poverty*. Washington DC: Economic Policy Institute.

Wolfers, Justin. 2015. "How Economists Came to Dominate the Conversation." in *New York Times*. New York, NY: The New York Times Company.

Wright, S. Paul, John T. White, William L. Sanders, and June C. Rivers. 2010. *SAS EVAAS Statistical Models*. Cary, NC: SAS Institute Inc. World Headquarters.

Table 1: Descriptive Statistics on Differences by School-Performance-Level in the District's Core Teachers of Grades 3-8

	2007-08			2012-13		
	Low performing schools	High performing schools	Diff[a]	Low performing schools	High performing schools	Diff[a]
Mean value-added score	-0.10	0.17	**160%**	-0.19	0.21	**189%**
	(0.85)	(0.86)		(1.02)	(0.98)	
College major:[b]						
Math	0.01	0.01	-0.01	0.01	0.01	0.00
Reading/Language arts	0.04	0.03	-0.02	0.04	0.04	0.00
Science	0.09	0.07	-0.02	0.06	0.06	0.00
Social studies	0.12	0.09	-0.03	0.13	0.11	-0.02
Education	0.15	0.17	0.02	0.23	0.25	0.01
Other	0.59	0.64	0.05	0.54	0.54	0.01
Master's degree or higher	0.35	0.34	-0.02	0.33	0.31	-0.01
Certification focus:[b]						
Math	0.14	0.10	-0.03	0.13	0.10	-0.02
Reading/Language arts	0.32	0.32	0.01	0.22	0.21	-0.01
Science	0.10	0.07	-0.02	0.07	0.07	0.00
Social studies	0.14	0.11	-0.03	0.10	0.09	-0.01
Special education	0.09	0.05	-0.04	0.07	0.05	-0.02
English proficiency	0.45	0.48	0.03	0.41	0.43	0.02
Only general	0.19	0.20	0.01	0.31	0.28	-0.03
Years of experience:						
3 or fewer	0.34	0.25	-0.09	0.29	0.22	-0.07
More than 3, up to 8	0.26	0.27	0.01	0.25	0.24	-0.02
More than 8, up to 15	0.17	0.20	0.03	0.24	0.26	0.02
More than 15	0.23	0.28	0.05	0.21	0.28	0.07
Male	0.27	0.21	-0.06	0.25	0.20	-0.05
Race:						
White	0.22	0.39	**0.17**	0.19	0.37	**0.19**
Black	0.48	0.29	**-0.19**	0.49	0.28	**-0.22**
Hispanic	0.19	0.22	0.03	0.22	0.26	0.03
Asian	0.05	0.04	-0.01	0.05	0.05	0.01
Other	0.06	0.06	0.00	0.05	0.04	-0.01
Teachers (n)	1,694	1,823		1,510	1,832	

Note: Standard deviations in parentheses below means.

a-To quantify differences by school-performance-level in teachers, a percent-change is used for continuous measures and a percentage-point-change for categorical measures. Bolded values represent relatively substantively larger differences (greater than 150% or 10 percentage points).

b-These categories are not mutually exclusive.

Table 2: Descriptive Statistics on Differences by School-Performance-Level in School Resources

	2007-08			2012-13		
	Low performing schools	High performing schools	Diff[a]	Low performing schools	High performing schools	Diff[a]
Per pupil expenditure	6425.77	6107.39	-5%	5570.71	5374.87	-4%
	(2706.03)	(2095.97)		(903.40)	(676.20)	
Mean salary for teachers with:						
1-5 years experience	43163.55	42528.68	-1%	46410.58	46359.76	0%
	(878.80)	(1447.37)		(608.86)	(685.33)	
20 years or more experience	59955.77	58199.50	-3%	63792.42	62607.92	-2%
	(1312.44)	(2358.61)		(2952.27)	(4453.86)	
Teacher-student ratio across whole school	0.07	0.06	-3%	0.06	0.06	0%
	(0.02)	(0.01)		(0.01)	(0.01)	
Teacher-student ratio across classrooms	0.06	0.05	**-26%**	0.04	0.03	**-25%**
	(0.10)	(0.01)		(0.03)	(0.02)	
Proportion core teachers with Master's or higher	0.35	0.32	-9%	0.30	0.28	-7%
	(0.15)	(0.11)		(0.10)	(0.10)	
All teachers' average years of experience	11.23	11.77	5%	10.43	11.67	**11%**
	(3.06)	(2.95)		(2.79)	(2.57)	
School has magnet program	0.28	0.48	**0.20**	0.30	0.54	**0.24**
Schools (n)	103	109		97	100	

Note: Standard deviations in parentheses below means. Standardized versions of the continuous measures in this table are used in multivariate analyses.

a-To quantify differences by school-performance-level in school resources, a percent-change is used for continuous measures and a percentage-point-change for categorical measures. Bolded values represent relatively substantively larger differences (greater than 150% or 10 percentage points).

Table 3: Descriptive Statistics on Differences by School-Performance-Level in the District's Students

	2007-08			2012-13		
	Low performing schools	High performing schools	Diff[a]	Low performing schools	High performing schools	Diff[a]
Average test score	-0.32	0.29	**211%**	-0.36	0.28	**229%**
	(0.74)	(0.90)		(0.76)	(0.88)	
Male	0.51	0.50	-0.01	0.50	0.51	0.00
Race:						
White	0.03	0.14	**0.11**	0.06	0.13	0.07
Black	0.35	0.22	**-0.13**	0.32	0.18	**-0.14**
Hispanic	0.61	0.59	-0.02	0.61	0.63	0.02
Asian	0.01	0.05	0.04	0.01	0.06	0.05
Other	0.01	0.01	0.00	0.01	0.01	0.00
Economic status:						
Not disadvantaged	0.08	0.29	**0.21**	0.07	0.27	**0.20**
Living in poverty	0.24	0.13	**-0.11**	0.50	0.33	**-0.17**
In English Learner program	0.27	0.25	-0.02	0.27	0.23	-0.04
In Gifted and Talented	0.07	0.22	**0.16**	0.10	0.28	**0.18**
In special education	0.13	0.09	-0.05	0.11	0.07	-0.04
Taking Spanish version of this year's test	0.08	0.11	0.03	0.08	0.08	0.00
Grade level:						
Three	0.16	0.21	0.05	0.17	0.20	0.03
Four	0.16	0.21	0.05	0.18	0.19	0.02
Five	0.15	0.20	0.05	0.17	0.18	0.01
Six	0.17	0.13	-0.04	0.16	0.15	-0.01
Seven	0.18	0.13	-0.06	0.16	0.15	-0.01
Eight	0.18	0.12	-0.05	0.16	0.13	-0.03
Students (n)	43,412	43,802		40,768	46,916	

Note: Standard deviations in parentheses below means. These estimates use student level data whereas main analyses use student level data linked to teachers for each subject, aggregated to subject-specific teacher level means, and then averaged across subjects.

a-To quantify differences by school-performance-level in students, a percent-change is used for continuous measures and a percentage-point-change for categorical measures. Bolded values represent relatively substantively larger differences (greater than 150% or 10 percentage points).

Table 4, Part 1 of 2: Contextual Factors that Mediate the Total Effect of Switching School-Performance-Level on Changes Over Time in Teachers' Value-Added Scores

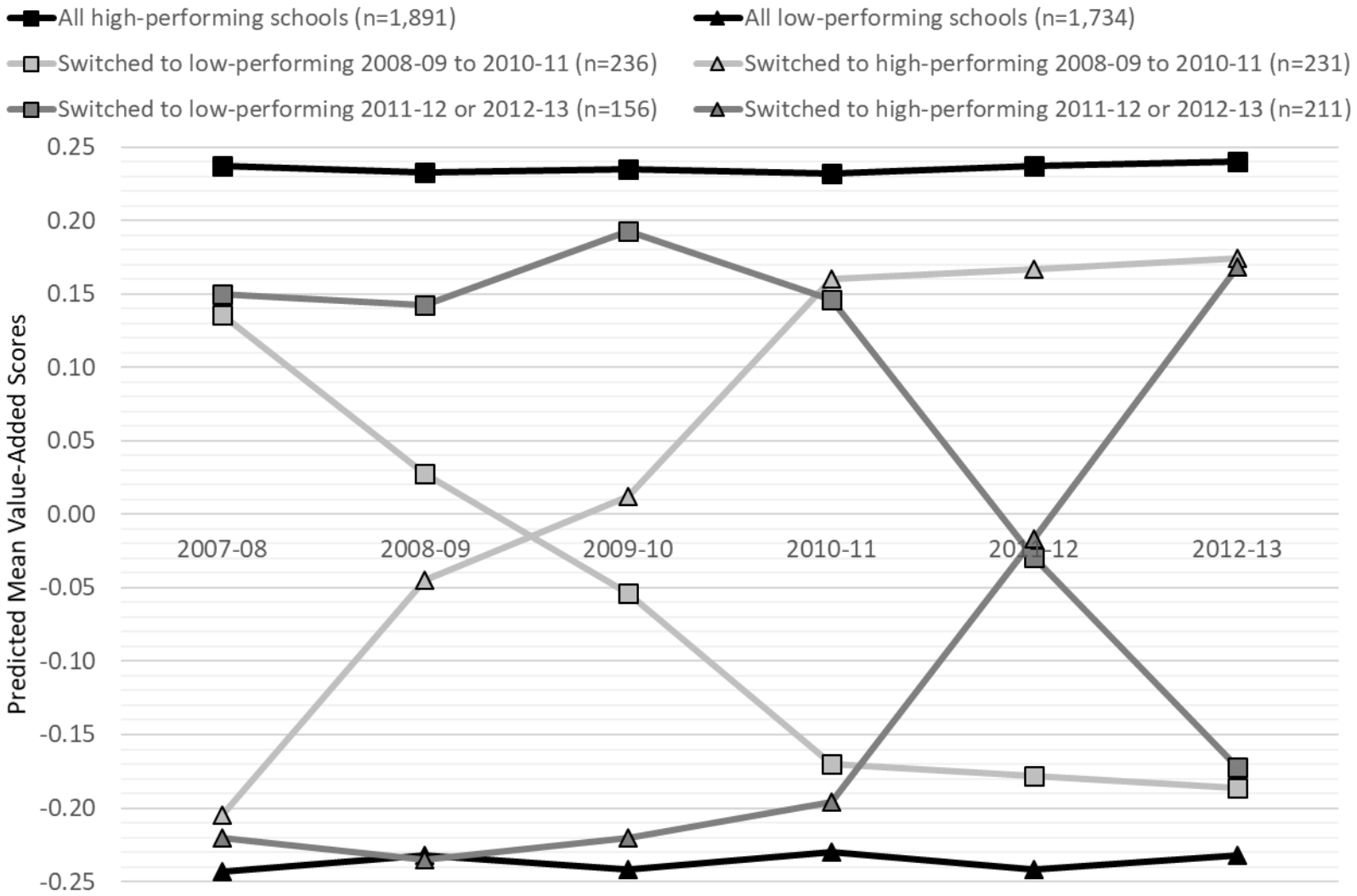
	Correlation with score change	Teachers initially in low-performing schools			Teachers initially in high-performing schools		
		Difference in means for teachers who:		Indirect effect of switching	Difference in means for teachers who:		Indirect effect of switching
		Stayed	Switched		Stayed	Switched	
Change in Value-Added Score		-0.04	0.38		-0.01	-0.29	
Changes in Characteristics of Teacher's Students							
Number of students	-0.06	0.02	0.07	-	0.03	0.00	-
Proportion male	-0.01	-0.12	0.19	1.0%	0.07	-0.05	0.5%
Proportion White	0.03	0.01	0.29	**3.2%**	-0.04	-0.11	-
Proportion Black	-0.03	0.03	-0.19	0.9%	0.01	0.07	**1.7%**
Proportion Asian	0.03	0.00	0.27	-	-0.04	-0.06	0.1%
Proportion not economically disadvantaged	0.08	0.07	**0.38**	**2.5%**	-0.12	-0.16	-
Proportion living in poverty	-0.04	0.05	**-0.41**	**2.1%**	-0.03	**0.27**	**2.2%**
Proportion in English Learner program	-0.04	-0.01	-0.08	0.9%	0.00	0.03	**1.4%**
Proportion in Gifted and Talented	0.04	-0.04	0.27	1.0%	0.01	-0.15	**1.3%**
Proportion in special education	-0.01	-0.09	-0.02	-	0.03	0.05	0.4%
Proportion taking Spanish version of test	-0.02	-0.01	-0.04	0.0%	-0.01	-0.03	0.0%
Proportion tested for grade 3	-0.03	-0.02	-0.02	0.4%	-0.05	-0.08	-
Proportion tested for grade 4	0.02	0.01	0.00	-	-0.01	0.02	0.6%
Proportion tested for grade 5	-0.01	0.00	0.02	1.3%	0.02	-0.01	-
Proportion tested for grade 6	-0.01	-0.02	0.03	**3.1%**	0.01	0.06	-
Proportion tested for grade 7	-0.03	-0.01	-0.03	-	0.01	0.00	0.2%
Proportion tested for grade 8	0.07	0.05	-0.01	-	0.03	0.00	**1.6%**

Table 4, Part 2 of 2: Contextual Factors that Mediate the Total Effect of Switching School-Performance-Level on Changes in Teachers' Value-Added Scores

	Correlation with score change	Teachers initially in low-performing schools			Teachers initially in high-performing schools		
		Mean for teachers who:		Indirect effect of switching	Mean for teachers who:		Indirect effect of switching
		Stayed	Switched		Stayed	Switched	
Changes in Characteristics of Students in Teacher's School							
Proportion not economically disadvantaged	0.04	-0.12	0.35	-	-0.02	**-0.45**	-
Proportion living in poverty	-0.03	0.15	-0.26	**7.0%**	0.10	**0.46**	-
Proportion White	0.02	0.00	0.36	-	-0.05	-0.13	-
Proportion Asian	0.03	0.00	0.26	-	-0.03	-0.07	0.9%
Proportion in English Learner program	-0.02	0.03	-0.17	-	-0.01	0.06	-
Proportion in Gifted and Talented	0.03	-0.04	**0.44**	-	0.05	-0.18	0.0%
Changes in School Resources							
Per pupil expenditures	0.06	0.01	-0.03	-	-0.02	0.00	-
Average salary: teachers with 1-5 years exp.	-0.03	-0.06	-0.32	0.5%	0.15	0.15	0.9%
Teacher-student ratio across whole school	0.02	0.02	0.04	-	0.04	0.04	0.0%
Teacher-student ratio across classrooms	0.03	0.00	-0.04	-	-0.02	-0.03	-
Proportion core teachers with Master's or higher	0.03	0.04	0.06	0.1%	0.05	0.05	-
Average years of experience for all teachers in school	-0.01	-0.01	0.21	-	0.16	0.05	-
School magnet status:[a]				-			-
Consistently in or not in magnet	-0.01	0.93	0.78		0.97	0.85	
Switched from a magnet to a non-magnet	-0.01	0.04	0.05		0.01	0.12	
Switched from a non-magnet to a magnet	0.15	0.03	0.17		0.02	0.03	

a-Mean change in value-added score is used instead of a correlation for school magnet status. Similarly, proportions are used to show changes in school magnet status for teachers who switch and stay.

Figure 1: Changes Over Time in Teachers' Average Value-Added Scores Depending on their School-Performance-Level-Trajectory



Note: Predicted means from a multilevel model with the intercept fixed at the teacher level, and the school-performance-level measure the only predictor. Means estimated for each category of the school-performance-level-trajectory indicator.

ONLINE APPENDIX A

Propensity score techniques may address selection bias better than standard regression techniques (Hong and Raudenbush 2005). Propensity scores are estimated from multilevel logistic regression models predicting switching, with a random-intercept at the level of each teacher's initial school (Online Table 6). A first model estimates the propensity to switch into a high-performing school among teachers whose initial school was low-performing. A second model estimates the propensity to switch into a low-performing school among teachers whose initial school was high-performing. These analyses do not differentiate by school year because exploratory analyses showed the results of switching are uniform regardless of the year of the switch. These models include measures from each teacher's first year in the district describing the teacher, their students, and their initial school. In cases of multicollinearity, the variables most highly correlated with changing school-performance-level are retained. Teachers are matched, within initial school-performance-level, using a nearest-neighbor matching technique developed by Robinson-Cimpian (2016). Matching success is assessed by comparing covariate balance, before and after matching, between teachers who switched school performance-levels and teachers who did not (Online Tables 7 and 8). Finally, multilevel linear models with random-intercepts at the level of the teacher's initial school predict: 1) changes in teachers' scores between their first and last years in the district to assess the robustness of the main results, and 2) teachers' final value-added score (their last year in the district before 2012-13) in order to understand whether score differences by school-performance-level are actually a product of how mobility relates to teachers' initial scores or years of experience (Online Tables 9 and 10). In other words, if teachers who switch into high-performing schools have higher initial value-added scores than teachers who switch into low-performing schools, the higher final score of teachers

in high-performing schools will be explained by these initial score differences rather than by changes in contextual factors. Similarly, if teachers who switch into high-performing schools are more inexperienced than teachers who switch into low-performing schools, higher-performing schools may have a larger share of teachers poised to experience improved scores with additional years of experience, regardless of the context. In this case, the higher final scores of teachers in high-performing schools will be explained by average differences in teacher experience. Each of these outcomes is predicted: 1) with teachers unmatched; 2) with teachers unmatched and baseline covariates included; 3) with teachers matched; 4) with teachers matched and propensity score included; 5) with teachers matched and baseline covariates included (to adjust for remaining differences between matched teachers (Robinson-Cimpian 2016)).

Online Table 1: Teachers and Students Included in Analyses and on Missing Data

	2007-08					2012-13				
Teachers with value-added score for any subject				3,770					3,643	
Excluded because:										
Not a core teacher of grades 3-8 per performance pay category[a]				71					153	
Charter school teacher				174					133	
Across-subject subtotal of teachers				3,525					3,357	
Average percent of teachers missing values across all variables				<1%					<1%	
	Mth	Rdg	LArts	Sci	SocSt	Mth	Rdg	LArts	Sci	SocSt
Subject-specific subtotals of teachers	1,809	1,904	1,993	1,174	1,269	1,651	1,689	1,778	1,078	1,133
Students in linkage files for grades 3-8				90,002					88,806	
Excluded because no values on any variables for any school years				2,788					1,122	
Across-subject subtotal of students				87,214					87,684	
	Mth	Rdg	LArts	Sci	SocSt	Mth	Rdg	LArts	Sci	SocSt
Excluded because:										
Ineligible grade level[b]	15	14	13	16,361	16,357	2	1	2	16,401	16,401
Not linked for subject	2,298	7,744	2,417	2,840	2,992	4,124	4,512	1,664	1,904	1,497
Only linked to teachers not in study	8,144	6,466	6,903	5,316	5,228	8,526	8,139	6,877	5,421	6,008
Subject-specific subtotals of students	76,757	72,990	77,881	62,697	62,637	75,032	75,032	79,141	63,958	63,778
Number of teachers linked to for each subject:										
One teacher	71,018	63,932	69,633	51,413	58,799	66,348	65,603	72,135	50,489	58,102
Two teachers	5,388	8,292	7,522	10,711	3,540	8,094	8,533	6,410	12,632	5,344
Three teachers	323	755	708	392	285	569	779	564	680	325
Four teachers	28	11	17	181	13	18	96	29	143	7
Five teachers	0	0	1	0	0	3	19	3	13	0
Six teachers	0	0	0	0	0	0	2	0	1	0
Linked students missing background data	4.1%	4.2%	4.1%	4.0%	3.9%	3.9%	3.8%	3.8%	4.0%	4.0%
Linked students missing test score data	5.0%	4.9%	5.1%	5.3%	5.2%	5.7%	5.6%	5.7%	6.1%	6.0%

Note: Mth=Math. Rdg=Reading. LArts=Language Arts. Sci=Science. SocSt=Social Studies.

a - Although only core teachers of grades 3-8 should have received value-added scores, some teachers' classifications were inconsistent.

b - A small number of students in grades 3-8 were tested for a grade level below grade 3 or above grade 8. No science or social studies tests were administered in grade 3.

Online Table 2, Part 1 of 2: Sensitivity Analysis - A More Nuanced Conceptualization of School Performance Level

Online Table 2-1: Coefficients from a Multilevel Linear Regression Model Predicting Changes Over Time in Teachers' Average Value-Added Scores with the Intercept Fixed at the Teacher Level

	B		(SE)
School performance level:			
Quartile 1 - Lowest Performing (ref)	-		
Quartile 2 - Low/Mid Performing	0.24	***	(0.03)
Quartile 3 - Mid/High Performing	0.46	***	(0.03)
Quartile 4 - Highest Performig	0.68	***	(0.04)
Constant	-0.37	***	(0.02)

Online Table 2-2: Contextual Factors that Mediate the Total Effect of Switching School-Performance-Level on Changes Over Time in Teachers' Value-Added Scores

	Switched two quartiles lower		Switched one quartile lower		Switched one quartile higher		Switched two quartiles higher	
	Mean change	Indirect effect	Mean change	Indirect effect	Mean change	Indirect effect	Mean change	Indirect effect
Change in Value-Added Score	-0.47		-0.17		0.32		0.58	
Changes in Characteristics of Teacher's Students								
Number of students	-0.05	-	-0.01	**3.2%**	-0.01	0.3%	0.21	-
Proportion male	-0.29	-	-0.07	0.7%	0.06	-	0.23	0.6%
Proportion White	-0.33	-	0.00	-	0.03	-	0.56	2.5%
Proportion Black	0.40	**17.6%**	0.02	-	-0.02	-	-0.34	-
Proportion Asian	-0.24	-	-0.01	0.0%	-0.01	0.0%	0.38	2.5%
Proportion tested for grade 3	-0.01	14.6%	-0.01	-	-0.01	-	0.05	**72.3%**
Proportion tested for grade 4	0.12	**90.3%**	0.00	2.8%	-0.01	-	-0.05	-
Proportion tested for grade 5	-0.17	-	-0.02	**10.0%**	0.00	**29.8%**	0.04	-
Proportion tested for grade 6	-0.01	**103.4%**	0.01	**9.9%**	0.03	**11.8%**	0.03	-
Proportion tested for grade 7	-0.09	**244.4%**	0.00	1.1%	-0.01	-	0.01	**64.2%**
Proportion tested for grade 8	0.15	-	0.02	-	0.00	-	-0.09	-
Proportion not economically disadvantaged	-0.43	15.2%	-0.03	-	0.02	1.0%	0.64	**21.0%**

Online Table 2, Part 2 of 2: Sensitivity Analysis - A More Nuanced Conceptualization of School Performance Level

Online Table 3-2, continued

	Switched two quartiles lower		Switched one quartile lower		Switched one quartile higher		Switched two quartiles higher	
Changes in Characteristics of Teacher's Students, continued	Mean change	Indirect effect	Mean change	Indirect effect	Mean change	Indirect effect	Mean change	Indirect effect
Proportion living in poverty	0.84	8.6%	0.06	**4.2%**	0.00	2.4%	-0.76	-
Proportion in English Learner program	0.04	10.1%	-0.02	1.3%	-0.02	-	-0.13	-
Proportion in Gifted and Talented	-0.22	3.7%	-0.04	1.6%	0.05	-	0.49	-
Proportion in special education	-0.08	0.1%	0.00	0.5%	0.03	1.3%	-0.12	-
Proportion taking Spanish version of test	0.01	-	-0.03	-	0.00	-	-0.02	-
Changes in Characteristics of Students in Teacher's School								
Proportion not economically disadvantaged	-0.36	6.9%	-0.04	2.0%	0.03	**2.9%**	0.73	-
Proportion living in poverty	0.81	-	0.14	-	0.01	-	-0.89	2.3%
Proportion White	-0.33	-	-0.02	-	0.05	0.0%	0.66	**31.8%**
Proportion Asian	-0.20	8.9%	-0.03	1.0%	0.01	0.0%	0.43	**12.3%**
Proportion in Gifted and Talented	-0.46	**27.1%**	-0.02	-	0.07	-	0.75	-
Proportion in English Learner program	0.05	1.1%	-0.01	1.0%	-0.02	0.5%	-0.26	4.5%
Changes in School Resources								
Per pupil expenditures	0.03	**17.6%**	-0.01	-	0.02	-	0.02	-
Average years of experience for all teachers in school	-0.08	0.6%	0.00	-	0.05	1.1%	0.23	-
Teacher-student ratio across classrooms	0.01	-	-0.01	-	0.01	-	-0.04	1.9%
Average salary: teachers with 1-5 years exp.	0.03	1.3%	0.01	-	-0.15	-	-0.15	-
Teacher-student ratio across whole school	0.03	15.8%	0.01	**5.8%**	0.02	**4.4%**	0.00	1.7%
Proportion core teachers with Master's or higher	-0.04	-	0.03	3.1%	0.06	**2.7%**	-0.07	-

a-Teachers who switched school performance level compared to teachers who stayed in school of same performance level. Analyses are conducted for each school year and then averaged. Bolded percentages represent the five mediators that contributed the most to the estimated effect of switching school performance levels on changes in teachers' value-added scores.

Online Table 3: Descriptive Statistics Comparing Teachers Included/Excluded in Analyses

	Included	Excluded	
		Mixed trajectory[a]	In district one year
Mean value-added score	-0.01	0.07	-0.22
Mean value-added score relative to school mean	-0.03	0.08	-0.19
College major:[b]			
Math	0.01	0.01	0.01
Reading/Language arts	0.04	0.04	0.03
Science	0.07	0.07	0.07
Social studies	0.12	0.14	0.09
Education	0.17	0.23	0.16
Other	0.59	0.52	0.65
Master's degree or higher	0.31	0.34	0.30
Certification focus:[b]			
Math	0.12	0.15	0.08
Reading/Language arts	0.27	0.25	0.21
Science	0.08	0.09	0.05
Social studies	0.11	0.10	0.07
Special education	0.08	0.12	0.05
English proficiency	0.48	0.38	0.62
Only general	0.22	0.29	0.21
Years of experience:			
3 or fewer	0.42	0.31	0.49
More than 3, up to 8	0.23	0.27	0.19
More than 8, up to 15	0.15	0.20	0.15
More than 15	0.20	0.23	0.17
Male	0.23	0.22	0.23
Race:			
White	0.31	0.18	0.36
Black	0.37	0.40	0.35
Hispanic	0.23	0.31	0.17
Asian	0.05	0.05	0.04
Other	0.04	0.06	0.07
Teachers (n)	4,459	467	1,455

Note: All measures from first year teacher in district between 2007-08 and 2012-13, except the change in value-added score is the difference between last and first years. Full details on exclusion criteria provided in Data and Methods section.

a-Teachers with mixed trajectories were too mobile to fit into one of the six school-performance-level-trajectories.

b-These categories are not mutually exclusive.

Online Table 4, Part 1 of 2: Sensitivity Analyses that Account for Time-Varying Characteristics of Teachers - Multilevel Linear Regression Models Predicting Teachers' Value-Added Scores Each School Year with Intercepts Fixed at the Teacher Level

	Model 1		Model 2		Model 3	
	Unadjusted Model		Adjusted by Previous Value-Added Score		Adjusted by Teacher Experience	
	B	(SE)	B	(SE)	B	(SE)
Contextual Factors: Teacher's Students						
Number of students	-0.13 ***	(0.01)	-0.15 ***	(0.02)	-0.15 ***	(0.02)
Proportion male	0.01	(0.01)	0.02 +	(0.01)	0.02 +	(0.01)
Proportion White	-0.02	(0.02)	-0.04	(0.03)	-0.05	(0.03)
Proportion Black	-0.07 **	(0.02)	-0.07 **	(0.03)	-0.09 **	(0.03)
Proportion Asian	0.02	(0.01)	0.02	(0.02)	0.02	(0.02)
Proportion not economically disadvantaged	0.08 ***	(0.01)	0.08 ***	(0.01)	0.08 ***	(0.01)
Proportion living in poverty	-0.01	(0.01)	-0.01	(0.02)	-0.02	(0.02)
Proportion in English Learner program	-0.01	(0.01)	-0.01	(0.02)	-0.02	(0.02)
Proportion in Gifted and Talented	-0.06 ***	(0.02)	-0.03	(0.02)	-0.03	(0.02)
Proportion in special education	0.03 *	(0.01)	0.05 **	(0.02)	0.05 **	(0.02)
Proportion taking Spanish version of test	-0.01	(0.01)	-0.01	(0.02)	-0.01	(0.02)
Proportion tested for grade 3	0.04 *	(0.02)	0.01	(0.02)	0.01	(0.02)
Proportion tested for grade 4	-0.06	(0.09)	0.01	(0.11)	0.03	(0.12)
Proportion tested for grade 5	-0.02	(0.09)	0.04	(0.11)	0.06	(0.11)
Proportion tested for grade 6	-0.04	(0.08)	0.01	(0.11)	0.02	(0.11)
Proportion tested for grade 7	0.06	(0.07)	0.13	(0.09)	0.13	(0.09)
Proportion tested for grade 8	0.06	(0.06)	0.12	(0.08)	0.12	(0.08)
	0.08	(0.06)	0.14 +	(0.08)	0.14 +	(0.08)
Contextual Factors: Characteristics of Students in Teacher's School						
Proportion not economically disadvantaged	-0.09	(0.06)	-0.21 **	(0.07)	-0.19 *	(0.07)
Proportion living in poverty	-0.05 +	(0.03)	-0.04	(0.03)	-0.01	(0.04)
Proportion White	0.04	(0.05)	0.13 *	(0.07)	0.13 +	(0.07)
Proportion Asian	0.02	(0.03)	0.04	(0.04)	0.03	(0.04)
Proportion in English Learner program	0.01	(0.03)	0.01	(0.04)	0.01	(0.04)
Proportion in Gifted and Talented	0.02	(0.03)	0.03	(0.04)	0.04	(0.04)

Online Table 4, Part 2 of 2: Sensitivity Analyses that Account for Time-Varying Characteristics of Teachers - Multilevel Linear Regression Models Predicting Teachers' Value-Added Scores Each School Year with Intercepts Fixed at the Teacher Level

	Model 1, cont. Unadjusted Model		Model 2, cont. Adjusted by Previous Value-Added Score		Model 3, cont. Adjusted by Teacher Experience	
	B	(SE)	B	(SE)	B	(SE)
Contextual Factors: School Resources						
Per pupil expenditures	0.14 ***	(0.04)	0.11 *	(0.05)	0.11 *	(0.05)
Average salary: teachers with 1-5 years experience	-0.03 **	(0.01)	-0.02 +	(0.01)	-0.02 +	(0.01)
Teacher-student ratio acros whole school	-0.02	(0.08)	0.05	(0.15)	0.07	(0.15)
Teacher-student ratio across classrooms	0.14 **	(0.05)	0.14 *	(0.07)	0.10	(0.07)
Proportion core teachers with Master's or higher	0.08 ***	(0.02)	0.08 ***	(0.02)	0.08 ***	(0.02)
Average years of experience for a	-0.05 **	(0.02)	-0.07 **	(0.02)	-0.06 *	(0.02)
teachers in school	0.06	(0.04)	0.18 **	(0.06)	0.21 **	(0.06)
School has magnet program						
Controls: Time Varying Characteristics of Teachers						
Last year's score relative to the school mean score			-0.14 ***	(0.01)		
Last year's experience relative to the district average					-0.02 **	(0.01)
Constant	0.00	(0.03)	-0.02	(0.04)	-0.04	(0.04)
Teachers(n)	6,381		4,721		4,648	

Note: Model 1 re-assesses main results but with no teachers excluded. Models 2 and 3 assess whether any relationship between context and teachers' value-added scores persists with controls for teachers' relative-score and relative-experience from the previous school year. These analyses include teachers excluded in main analyses (teachers in district one year and teachers with mixed school-performance-level-trajectories), although Models 2 and 3 exclude teachers missing on the control variables becaues they were in the district discontinuously.

Online Table 5, Part 1 of 2: Sensitivity Analyses that Include all Teachers - Teacher and Contextual Factors that Mediate the Effect of Teaching in a Low-Performing School on Teachers' Value-Added Scores in Each School Year

	Indirect Effects[a]					
	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13
Characteristics of Teacher's Students						
Number of students	1.6%	3.9%	0.8%	1.4%	1.0%	0.4%
Proportion male	0.5%	-	1.2%	**8.2%**	2.7%	0.9%
Proportion Black	-	-	-	6.5%	2.6%	2.8%
Proportion tested for grade 4	2.2%	2.1%	1.1%	-	4.1%	1.2%
Proportion tested for grade 5	3.1%	3.4%	2.8%	1.2%	2.9%	0.5%
Proportion tested for grade 6	-	-	-	-	-	-
Proportion tested for grade 7	-	-	-	-	0.5%	-
Proportion tested for grade 8	-	-	-	-	-	-
Proportion not economically disadvantaged	**39.6%**	**108.9%**	**39.5%**	**10.8%**	**15.6%**	-2.7%
Proportion eligible for free lunch	-	-	-	**8.4%**	**5.3%**	-
Proportion in Gifted and Talented	-	-	0.3%	4.0%	**21.3%**	**17.9%**
Proportion in special education	2.9%	**7.1%**	5.4%	1.6%	-	1.7%
Proportion taking Spanish version of test	**4.7%**	2.0%	1.0%	-	-	-
Characteristics of Students in Teacher's School						
Proportion not economically disadvantaged	-	-	-	-	-	-
Proportion eligible for free lunch	**28.7%**	**18.2%**	**18.7%**	**18.2%**	-	**36.5%**
Proportion Asian	-	-	**18.1%**	1.2%	**14.6%**	1.6%
Proportion in Gifted and Talented	**14.8%**	**5.9%**	**8.9%**	2.9%	-	-
Proportion in special education	**18.7%**	-	**47.4%**	**10.1%**	-	**20.2%**
Proportion in English Learner program	-	**5.2%**	1.6%	1.8%	-	-
School Resources						
Per pupil expenditures	-	-	-	-	2.3%	-
Average years of experience across all teachers	-	-	-	-	-	-
Teacher-student ratio across classrooms	3.0%	-	0.7%	0.3%	0.7%	**14.7%**
Average salary for teachers with 20 years or more experience	-	-	-	1.7%	**11.8%**	-
Proportion core teachers with Master's or higher	-	0.4%	7.2%	-	-	1.3%
School magnet status	-	-	-	-	-	-

Online Table 5, Part 2 of 2: Sensitivity Analyses that Include all Teachers - Teacher and Contextual Factors that Mediate the Effect of Teaching in a Low-Performing School on Teachers' Value-Added Scores in Each School Year

	Indirect Effects,[a] continued					
	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13
Characteristics of Teachers						
Years of experience	0.9%	-	-	0.1%	0.2%	**3.7%**
College major:[b]						
Math	-	-	-	0.1%	0.0%	0.1%
Reading/Language arts	0.3%	0.1%	0.1%	0.0%	-	0.1%
Science	-	0.4%	0.0%	-	0.0%	0.0%
Social studies	-	-	-	-	-	0.3%
Education	-	-	-	-	-	-
Master's degree or higher	0.1%	0.0%	0.1%	0.1%	0.0%	-
Certification focus:[b]						
Math	0.4%	2.4%	1.5%	1.3%	-	-
Reading/Language arts	0.1%	-	0.0%	0.1%	0.1%	0.1%
Science	0.2%	-	0.4%	-	-	0.0%
Social studies	0.7%	0.4%	-	1.3%	0.3%	0.2%
Special education	1.8%	0.2%	0.7%	2.7%	-	0.6%
English proficiency	-	-	0.0%	0.2%	0.2%	0.2%
Male	2.6%	2.2%	1.6%	1.5%	0.2%	0.9%
Race	-	0.7%	-	-	-	-
Teachers (n)	3,517	3,331	3,618	3,526	3,349	3,342

Note: These sensitivity analyses use the same technique as results in Table 4 but include all teachers (which forces a non-longitudinal approach). With this technique based in regression modeling, models do not support collinear potential-mediators. These analyses include the potential-mediator from each collinear pair that related more closely to teachers' value-added scores in exploratory analyses.

a-Indirect effects indicate the percent of the total effect of teaching in a low- rather than high-performance school that is mediated through each teacher, school, and student characteristic. A hyphen indicates the measure did not function as a mediator. The five largest mediators in each school year are bolded.

b-These categories are not mutually exclusive.

Online Table 6, Part 1 of 3: Multilevel Logistic Regression Models Predicting Teachers Switch School-Performance-Levels, with Random-Intercept at Origin-School Level

	Model 1: Switch from low- to high-performing[a]		Model 2: Switch from high- to low-performing[b]	
	B	(SE)	B	(SE)
Characteristics of Teachers				
Mean value-added score relative to school mean	0.20	(0.09)	0.19	(0.11)
Years of experience:				
3 or fewer	-0.02	(0.23)	-0.82	(0.26)
More than 3, up to 8 (ref)	-		-	
More than 8, up to 15	-0.01	(0.24)	-0.56	(0.26)
More than 15	-0.48	(0.28)	-0.88	(0.29)
Degree:[c]				
Math	0.29	(0.56)	-0.16	(1.04)
Reading/Language arts	0.01	(0.37)	-0.53	(0.55)
Science	0.04	(0.29)	-0.13	(0.35)
Social studies	0.08	(0.22)	0.20	(0.29)
Education	-0.09	(0.22)	-0.24	(0.24)
Master's degree or higher	0.08	(0.17)	0.13	(0.20)
Certification focus:[c]				
Math	-0.23	(0.25)	0.27	(0.33)
Reading/Language arts	-0.19	(0.20)	-0.15	(0.25)
Science	0.26	(0.27)	0.06	(0.38)
Social studies	-0.75	(0.30)	-0.13	(0.36)
Special education	-0.05	(0.30)	0.10	(0.41)
English proficiency	0.16	(0.18)	-0.04	(0.22)
Male	-0.15	(0.17)	0.11	(0.21)
Race:				
White, non-Hispanic (ref)	-		-	
Black, non-Hispanic	-0.04	(0.21)	1.20	(0.27)
Hispanic, white	0.59	(0.24)	0.49	(0.29)
Asian, non-Hispanic	0.37	(0.35)	0.84	(0.44)
Other	0.43	(0.36)	1.27	(0.40)

Online Table 6, Part 2 of 3: Multilevel Logistic Regression Models Predicting Teachers Switch School-Performance-Levels, with Random-Intercept at School Level

	Model 1, continuted: Switch from low- to high-performing[a]		Model 2, continued: Switch from high- to low-performing[b]	
	B	(SE)	B	(SE)
Characteristics of Teacher's Students				
Number of students	0.01	(0.11)	0.16	(0.16)
Proportion male	-0.06	(0.07)	0.09	(0.09)
Proportion white	**0.54**	(0.47)	#	
Proportion black	#		**0.15**	(0.20)
Proportion Hispanic	-0.29	(0.23)	#	
Proportion Asian	-0.06	(0.13)	-0.02	(0.08)
Proportion other race	#		#	
Proportion not economically disadvantaged	-0.13	(0.31)	-0.08	(0.19)
Proportion eligible for reduced lunch	**0.07**	(0.11)	-0.05	(0.10)
Proportion eligible for free lunch	-0.11	(0.17)	#	
Proportion living in poverty	#		#	
Proportion in English Learner program	-0.02	(0.14)	**0.02**	(0.14)
Proportion in Gifted and Talented	**0.06**	(0.16)	-0.08	(0.12)
Proportion in special education	-0.12	(0.09)	-0.10	(0.13)
Proportion taking Spanish version of test	-0.15	(0.12)	#	
Proportion tested for grade 3	0.18	(0.98)	1.66	(1.61)
Proportion tested for grade 4	0.26	(0.98)	1.69	(1.61)
Proportion tested for grade 5	0.31	(0.94)	1.60	(1.54)
Proportion tested for grade 6	-0.28	(0.79)	1.29	(1.30)
Proportion tested for grade 7	-0.21	(0.73)	1.32	(1.21)
Proportion tested for grade 8	-0.33	(0.75)	1.46	(1.24)
Characteristics of Students in Teacher's School				
Number of students	0.03	(0.36)	-1.12	(0.41)
Proportion not economically disadvantaged	**0.64**	(0.85)	#	
Proportion eligible for reduced lunch	-0.04	(0.28)	-0.18	(0.23)
Proportion eligible for free lunch	#		#	
Proportion living in poverty	-0.26	(0.35)	**0.72**	(0.35)
Proportion white	**1.78**	(1.11)	#	
Proportion black	-0.07	(0.36)	#	
Proportion Hispanic	#		#	
Proportion Asian	0.55	(0.48)	-0.98	(0.36)
Proportion in English Learner program	-0.10	(0.39)	**0.18**	(0.31)
Proportion in Gifted and Talented	**0.91**	(0.52)	-0.41	(0.34)
Proportion in special education	**0.52**	(0.50)	**1.63**	(0.72)

Online Table 6, Part 3 of 3: Multilevel Logistic Regression Models Predicting Teachers Switch School-Performance-Levels, with Random-Intercept at School Level

	Model 1, continuted: Switch from low- to high-performing[a]		Model 2, continued: Switch from high- to low-performing[b]	
	B	(SE)	B	(SE)
School Resources				
Per pupil expenditures	-0.97	(0.67)	-0.55	(0.73)
Average salary for teachers with 1-5 years experience	#		#	
Average salary for teachers with 20 years or more experience	-0.24	(0.18)	-0.14	(0.21)
Teacher-student ratio across whole school	#		#	
Teacher-student ratio across classrooms	#		#	
Proportion core teachers with Master's or higher	-0.42	(0.20)	-0.19	(0.25)
Average years of experience across all teachers	0.13	(0.41)	#	
Average years of experience across core teachers	-0.13	(0.37)	0.07	(0.24)
School is a magnet	0.17	(0.50)	-0.53	(0.52)
Constant	0.04	(0.61)	-1.73	(0.46)

Note: Propensity scores were estimated from these models. All measures from teacher's first year in the district between 2007-08 and 2012-13.

a-This model predicts switching into high-performing schools in any year (n=442), among the 2,176 teachers who started in low-performing schools.

b-This model predicts switching into low-performing schools in any year (n=392), among the 2,283 teachers who started in high-performing schools.

c-These categories are not mutually exclusive, such that each reference category includes teachers without the specific degree or certification.

#-This measure was excluded because it was multicollinear with other predictor(s) but less highly correlated with switching school-performance-level.

Online Table 7, Part 1 of 3: Covariate Balance Before and After Propensity-Score Matching for Teachers Who Started in Low-Performing Schools

	Unmatched			Matched		
	Teachers who:			Teachers who:		
Characteristics of Teachers	Stayed	Switched	P-value	Stayed	Switched	P-value
Propensity score				-1.23	-1.23	1.000
Mean value-added score relative to school mean	-0.03	0.05	0.039	0.01	0.04	0.650
Years of experience:						
3 or fewer	0.46	0.50	0.165	0.51	0.50	0.808
More than 3, up to 8	0.22	0.24	0.246	0.23	0.24	0.572
More than 8, up to 15	0.13	0.13	0.963	0.15	0.14	0.548
More than 15	0.19	0.12	0.002	0.12	0.12	0.773
Degree:[a]						
Math	0.02	0.02	0.897	0.02	0.02	0.509
Reading/Language arts	0.05	0.04	0.549	0.04	0.04	0.742
Science	0.08	0.09	0.375	0.09	0.08	0.518
Social studies	0.14	0.16	0.274	0.16	0.16	0.803
Education	0.16	0.15	0.632	0.12	0.15	0.278
Master's degree or higher	0.32	0.29	0.254	0.29	0.29	0.946
Certification focus:[a]						
Math	0.15	0.10	0.005	0.12	0.10	0.477
Reading/Language arts	0.30	0.21	0.000	0.19	0.21	0.474
Science	0.10	0.08	0.135	0.09	0.08	0.577
Social studies	0.14	0.05	0.000	0.07	0.05	0.490
Special education	0.11	0.06	0.001	0.09	0.06	0.230
English proficiency	0.45	0.55	0.000	0.57	0.55	0.496
Male	0.27	0.24	0.322	0.28	0.25	0.361
Race:						
White, non-Hispanic	0.23	0.25	0.451	0.19	0.24	0.116
Black, non-Hispanic	0.48	0.35	0.000	0.34	0.36	0.433
Hispanic, white	0.20	0.29	0.000	0.36	0.29	0.085
Asian, non-Hispanic	0.05	0.05	0.541	0.07	0.06	0.414
Other	0.04	0.05	0.207	0.05	0.05	0.894

Online Table 7, Part 2 of 3: Covariate Balance Before and After Propensity-Score Matching for Teachers Who Started in Low-Performing Schools

	Unmatched			Matched		
	Teachers who:			Teachers who:		
	Stayed	Switched	P-value	Stayed	Switched	P-value
Characteristics of Teacher's Students						
Number of students	0.15	-0.14	0.000	0.03	-0.12	0.030
Proportion male	-0.09	-0.10	0.852	-0.14	-0.09	0.607
Proportion white	-0.41	-0.37	0.000	-0.39	-0.39	0.824
Proportion black	#					
Proportion Hispanic	#					
Proportion Asian	-0.29	-0.23	0.010	-0.29	-0.24	0.063
Proportion other race	-0.11	-0.02	0.015	-0.11	-0.05	0.299
Proportion not economically disadvantaged	-0.41	-0.38	0.034	-0.42	-0.40	0.442
Proportion eligible for reduced lunch	-0.27	-0.09	0.000	-0.20	-0.11	0.165
Proportion eligible for free lunch	0.26	0.37	0.002	0.42	0.39	0.552
Proportion living in poverty	#					
Proportion in English Learner program	-0.01	0.15	0.001	0.21	0.14	0.377
Proportion in Gifted and Talented	-0.40	-0.29	0.000	-0.34	-0.29	0.220
Proportion in special education	0.28	-0.03	0.000	0.03	-0.03	0.330
Proportion taking Spanish version of test	-0.09	0.04	0.011	-0.02	0.01	0.764
Proportion tested for grade 3	-0.15	-0.03	0.012	-0.09	-0.04	0.474
Proportion tested for grade 4	-0.13	0.05	0.000	0.03	0.05	0.848
Proportion tested for grade 5	-0.17	0.06	0.000	-0.02	0.04	0.385
Proportion tested for grade 6	0.18	-0.08	0.000	-0.02	-0.07	0.498
Proportion tested for grade 7	0.19	0.03	0.007	0.15	0.05	0.203
Proportion tested for grade 8	0.18	-0.05	0.000	-0.03	-0.04	0.908

Online Table 7, Part 3 of 3: Covariate Balance Before and After Propensity-Score Matching for Teachers Who Started in Low-Performing Schools

	Unmatched			Matched		
	Teachers who:			Teachers who:		
	Stayed	Switched	P-value	Stayed	Switched	P-value
Characteristics of Students in Teacher's School						
Number of students	0.19	0.17	0.636	0.17	0.18	0.800
Proportion not economically disadvantaged	-0.50	-0.51	0.483	-0.54	-0.54	0.935
Proportion eligible for reduced lunch	-0.32	-0.15	0.000	-0.24	-0.18	0.265
Proportion eligible for free lunch	#					
Proportion living in poverty	0.50	0.27	0.000	0.32	0.29	0.473
Proportion white	-0.41	-0.37	0.000	-0.39	-0.40	0.100
Proportion black	0.07	-0.20	0.000	-0.29	-0.17	0.079
Proportion Hispanic	#					
Proportion Asian	-0.30	-0.22	0.000	-0.30	-0.22	0.003
Proportion in Gifted and Talented	-0.34	-0.27	0.000	-0.28	-0.27	0.488
Proportion in special education	0.28	0.06	0.000	0.12	0.06	0.163
Proportion in English Learner program	0.12	0.45	0.000	0.45	0.44	0.913
School Resources						
Per pupil expenditures	-0.01	-0.07	0.000	-0.07	-0.07	0.746
Average salary for teachers with 1-5 years experience	#					
Average salary for teachers with 20 years or more experience	0.35	0.28	0.014	0.21	0.32	0.015
Teacher-student ratio across whole school	#					
Teacher-student ratio across classrooms	#					
Proportion core teachers with Master's or higher	0.02	-0.29	0.000	-0.31	-0.30	0.853
Average years of experience across all teachers	-0.23	-0.35	0.010	-0.37	-0.36	0.885
Average years of experience across core teachers	-0.27	-0.36	0.068	-0.38	-0.37	0.952
School is a magnet	0.43	0.36	0.014	0.34	0.35	0.797

a-These categories are not mutually exclusive.

#-Indicator excluded because multicollinear with other indicator(s) and less correlated with switching school-performance-level.

Online Table 8, Part 1 of 3: Covariate Balance Before and After Propensity-Score Matching for Teachers Who Started in High-Performing Schools

	Unmatched			Matched		
	Teachers who:			Teachers who:		
	Stayed	Switched	P-value	Stayed	Switched	P-value
Characteristics of Teachers						
Propensity score				-1.09	-1.09	0.999
Mean value-added score relative to school mean	-0.05	-0.03	0.619	-0.11	-0.04	0.302
Years of experience:						
3 or fewer	0.38	0.33	0.062	0.39	0.36	0.422
More than 3, up to 8	0.23	0.26	0.273	0.26	0.24	0.678
More than 8, up to 15	0.16	0.21	0.016	0.18	0.18	0.931
More than 15	0.23	0.20	0.273	0.17	0.22	0.128
Degree:[a]						
Math	0.01	0.01	0.441	0.00	0.01	0.564
Reading/Language arts	0.04	0.02	0.008	0.02	0.02	0.762
Science	0.07	0.06	0.497	0.08	0.07	0.629
Social studies	0.11	0.09	0.521	0.15	0.09	0.072
Education	0.19	0.16	0.254	0.17	0.15	0.713
Master's degree or higher	0.30	0.32	0.472	0.30	0.32	0.579
Certification focus:[a]						
Math	0.11	0.07	0.053	0.07	0.07	0.773
Reading/Language arts	0.28	0.21	0.013	0.14	0.22	0.023
Science	0.08	0.04	0.023	0.04	0.04	1.000
Social studies	0.10	0.06	0.004	0.06	0.06	1.000
Special education	0.06	0.05	0.274	0.07	0.05	0.496
English proficiency	0.47	0.53	0.016	0.62	0.55	0.146
Male	0.20	0.22	0.357	0.24	0.23	0.665
Race:						
White, non-Hispanic	0.43	0.18	0.000	0.16	0.20	0.160
Black, non-Hispanic	0.26	0.42	0.000	0.35	0.37	0.639
Hispanic, white	0.23	0.30	0.003	0.35	0.33	0.643
Asian, non-Hispanic	0.05	0.04	0.890	0.05	0.04	0.780
Other	0.04	0.06	0.034	0.10	0.06	0.172

Online Table 8, Part 2 of 3: Covariate Balance Before and After Propensity-Score Matching for Teachers Who Started in High-Performing Schools

	Unmatched			Matched		
	Teachers who:			Teachers who:		
	Stayed	Switched	P-value	Stayed	Switched	P-value
Characteristics of Teacher's Students						
Number of students	-0.02	-0.45	0.000	-0.37	-0.41	0.562
Proportion male	-0.02	-0.04	0.709	0.03	-0.01	0.711
Proportion white	#					
Proportion black	-0.23	-0.01	0.000	-0.27	-0.14	0.127
Proportion Hispanic	#					
Proportion Asian	#					
Proportion other race	0.12	-0.03	0.032	-0.15	-0.02	0.041
Proportion not economically disadvantaged	0.50	-0.23	0.000	-0.20	-0.22	0.728
Proportion eligible for reduced lunch	0.26	0.03	0.000	0.08	0.08	0.953
Proportion eligible for free lunch	#					
Proportion living in poverty	#					
Proportion in English Learner program	-0.11	0.16	0.000	0.46	0.22	0.021
Proportion in Gifted and Talented	0.39	-0.16	0.000	-0.18	-0.14	0.525
Proportion in special education	-0.14	-0.08	0.282	-0.10	-0.09	0.927
Proportion taking Spanish version of test	#					
Proportion tested for grade 3	0.11	0.32	0.001	0.23	0.33	0.337
Proportion tested for grade 4	0.03	0.21	0.002	0.22	0.15	0.493
Proportion tested for grade 5	0.00	0.14	0.014	0.22	0.14	0.473
Proportion tested for grade 6	-0.05	-0.33	0.000	-0.26	-0.31	0.345
Proportion tested for grade 7	-0.06	-0.26	0.000	-0.27	-0.24	0.612
Proportion tested for grade 8	-0.07	-0.24	0.001	-0.29	-0.21	0.094

Online Table 8, Part 3 of 3: Covariate Balance Before and After Propensity-Score Matching for Teachers Who Started in High-Performing Schools

	Unmatched			Matched		
	Teachers who:			Teachers who:		
	Stayed	Switched	P-value	Stayed	Switched	P-value
Characteristics of Students in Teacher's School						
Number of students	0.47	0.05	0.000	0.09	0.10	0.847
Proportion not economically disadvantaged	#					
Proportion eligible for reduced lunch	0.43	0.20	0.000	0.33	0.25	0.206
Proportion eligible for free lunch	#					
Proportion living in poverty	-0.63	0.09	0.000	-0.07	-0.02	0.322
Proportion white	#					
Proportion black	#					
Proportion Hispanic	#					
Proportion Asian	0.51	-0.25	0.000	-0.26	-0.23	0.313
Proportion in Gifted and Talented	0.68	-0.07	0.000	0.02	-0.04	0.214
Proportion in special education	-0.17	-0.17	0.783	-0.23	-0.20	0.107
Proportion in English Learner program	0.00	0.53	0.000	0.65	0.59	0.402
School Resources						
Per pupil expenditures	-0.09	-0.10	0.500	-0.12	-0.10	0.149
Average salary for teachers with 1-5 years experience	#					
Average salary for teachers with 20 years or more experience	-0.60	-0.09	0.000	-0.10	-0.13	0.676
Teacher-student ratio across whole school	#					
Teacher-student ratio across classrooms	#					
Proportion core teachers with Master's or higher	-0.17	-0.28	0.017	-0.48	-0.33	0.039
Average years of experience across all teachers	#					
Average years of experience across core teachers	0.07	0.12	0.192	-0.05	0.12	0.030
School is a magnet	0.65	0.33	0.000	0.35	0.38	0.577

a-These categories are not mutually exclusive.

#-Indicator excluded because multicollinear with other indicator(s) and less correlated with switching school-performance-level.

Online Table 9, Part 1 of 4: Propensity Score Estimates of Effect of Switching from Low- to High-Performing School on Teachers' Value-Added Scores

	Change in Value-Added Score[a]				Final Value-Added Score[b]			
	Teachers Unmatched		Teachers Matched		Teachers Unmatched		Teachers Matched	
	B	(SE)	B	(SE)	B	(SE)	B	(SE)
	Unadjusted Models							
Characteristics of Teachers	Model A1		Model A3		Model B1		Model B3	
Switched from low- to high-performing school	0.41 ***	(0.06)	0.46 ***	(0.09)	0.50 ***	(0.05)	0.50 ***	(0.07)
Constant	-0.04	(0.03)	-0.10	(0.06)	-0.21 ***	(0.03)	-0.21 ***	(0.05)
	Models Adjusted by Propensity Scores							
			Model A4				Model B4	
Switched from low- to high-performing school			0.46 ***	(0.09)			0.49 ***	(0.07)
Propensity score			-0.01	(0.04)			0.15 ***	(0.04)
Constant			-0.11	(0.09)			-0.02	(0.07)
	Models Adjusted by Baseline Covariates							
	Model A2		Model A5		Model B2		Model A5	
Switched from low- to high-performing school	0.47 ***	(0.06)	0.50 ***	(0.07)	0.45 ***	(0.05)	0.46 ***	(0.06)
Mean value-added score relative to school mean	-0.73 ***	(0.03)	-0.70 ***	(0.05)	0.29 ***	(0.03)	0.32 ***	(0.05)
Years of experience:								
3 or fewer (ref)	-		-		-		-	
More than 3, up to 8	-0.14 *	(0.06)	-0.10	(0.10)	-0.13 *	(0.05)	-0.09	(0.09)
More than 8, up to 15	-0.07	(0.07)	0.04	(0.14)	-0.07	(0.07)	0.04	(0.13)
More than 15	-0.09	(0.07)	-0.13	(0.13)	-0.09	(0.07)	-0.09	(0.12)
Degree:[a]								
Math	-0.01	(0.18)	-0.25	(0.28)	0.05	(0.17)	-0.15	(0.28)
Reading/Language arts	-0.28 **	(0.11)	-0.17	(0.15)	-0.25 *	(0.10)	-0.15	(0.14)
Science	0.07	(0.08)	0.07	(0.15)	0.04	(0.08)	0.09	(0.14)
Social studies	0.03	(0.07)	0.08	(0.11)	0.04	(0.06)	0.10	(0.10)
Education	-0.12 *	(0.06)	-0.07	(0.10)	-0.13 *	(0.06)	-0.08	(0.10)

Online Table 9, Part 2 of 4: Propensity Score Estimates of Effect of Switching from Low- to High-Performing School on Teachers' Value-Added Scores

	Change in Value-Added Score[a]				Final Value-Added Score[b]			
	Teachers Unmatched		Teachers Matched		Teachers Unmatched		Teachers Matched	
	Models Adjusted by Baseline Covariates, continued							
	Model A2		Model A5		Model B2		Model A5	
Characteristics of Teachers, continued	B	(SE)	B	(SE)	B	(SE)	B	(SE)
Master's degree or higher	-0.06	(0.05)	-0.03	(0.08)	-0.06	(0.05)	-0.09	(0.07)
Certification focus:[a]								
Math	0.07	(0.07)	0.00	(0.13)	0.03	(0.07)	-0.04	(0.14)
Reading/Language arts	-0.03	(0.06)	0.01	(0.10)	-0.03	(0.05)	-0.01	(0.10)
Science	0.03	(0.08)	0.05	(0.13)	0.07	(0.07)	0.02	(0.14)
Social studies	-0.06	(0.07)	-0.25	(0.18)	-0.04	(0.07)	-0.30 +	(0.18)
Special education	-0.08	(0.08)	-0.09	(0.12)	-0.03	(0.08)	-0.02	(0.11)
English proficiency	-0.06	(0.05)	-0.18 *	(0.09)	-0.04	(0.05)	-0.19 *	(0.08)
Male	-0.07	(0.05)	-0.03	(0.07)	-0.09 +	(0.05)	-0.06	(0.07)
Race:								
White, non-Hispanic (ref)	-		-		-		-	
Black, non-Hispanic	0.07	(0.06)	-0.02	(0.10)	0.08	(0.06)	-0.01	(0.10)
Hispanic, white	0.09	(0.07)	-0.02	(0.10)	0.06	(0.07)	-0.02	(0.10)
Asian, non-Hispanic	0.28 **	(0.10)	0.47 +	(0.27)	0.31 **	(0.10)	0.48 *	(0.24)
Other	-0.05	(0.11)	-0.04	(0.23)	-0.01	(0.11)	0.00	(0.21)
Characteristics of Teacher's Students								
Number of students	0.04	(0.03)	0.04	(0.04)	0.04	(0.03)	0.01	(0.04)
Proportion male	0.03	(0.02)	0.02	(0.03)	0.02	(0.02)	0.02	(0.03)
Proportion white	-0.17	(0.14)	-0.23	(0.22)	-0.19	(0.13)	-0.22	(0.22)
Proportion Asian	0.01	(0.07)	-0.11	(0.11)	-0.01	(0.07)	-0.17	(0.11)
Proportion other race	0.03	(0.03)	0.04	(0.04)	0.01	(0.03)	0.00	(0.04)

Online Table 9, Part 3 of 4: Propensity Score Estimates of Effect of Switching from Low- to High-Performing School on Teachers' Value-Added Scores

	Change in Value-Added Score[a]				Final Value-Added Score[b]			
	Teachers Unmatched		Teachers Matched		Teachers Unmatched		Teachers Matched	
	Models Adjusted by Baseline Covariates, continued							
	Model A2		Model A5		Model B2		Model A5	
Characteristics of Teacher's Students, continued	B	(SE)	B	(SE)	B	(SE)	B	(SE)
Proportion not economically disadvantaged	0.21 *	(0.09)	0.13	(0.13)	0.15 +	(0.08)	0.03	(0.12)
Proportion eligible for reduced lunch	0.00	(0.03)	0.01	(0.04)	0.01	(0.03)	0.01	(0.04)
Proportion eligible for free lunch	-0.04	(0.05)	-0.11	(0.07)	-0.04	(0.05)	-0.15 *	(0.07)
Proportion in English Learner program	-0.03	(0.04)	0.05	(0.05)	-0.04	(0.04)	0.04	(0.05)
Proportion in Gifted and Talented	0.01	(0.05)	0.04	(0.08)	0.03	(0.05)	0.07	(0.08)
Proportion in special education	0.01	(0.02)	-0.01	(0.04)	0.01	(0.02)	-0.01	(0.04)
Proportion taking Spanish version of test	0.01	(0.04)	-0.01	(0.05)	0.01	(0.03)	0.01	(0.05)
Proportion tested for grade 3	-0.58 *	(0.26)	-1.11	(0.75)	-0.48 +	(0.25)	-0.86	(0.70)
Proportion tested for grade 4	-0.59 *	(0.26)	-1.13	(0.75)	-0.48 +	(0.25)	-0.85	(0.70)
Proportion tested for grade 5	-0.54 *	(0.25)	-1.02	(0.72)	-0.43 +	(0.24)	-0.77	(0.67)
Proportion tested for grade 6	-0.50 *	(0.21)	-1.03 +	(0.62)	-0.44 *	(0.21)	-0.80	(0.58)
Proportion tested for grade 7	-0.43 *	(0.20)	-0.96 +	(0.57)	-0.39 *	(0.19)	-0.76	(0.53)
Proportion tested for grade 8	-0.38 +	(0.20)	-0.96	(0.60)	-0.34 +	(0.20)	-0.75	(0.55)
Characteristics of Students in Teacher's School								
Number of students	0.14 *	(0.06)	0.29 **	(0.11)	0.20 ***	(0.05)	0.32 ***	(0.08)
Proportion not economically disadvantaged	-0.07	(0.18)	-0.55	(0.37)	-0.07	(0.16)	-0.37	(0.27)
Proportion eligible for reduced lunch	-0.13 *	(0.06)	-0.25 **	(0.09)	-0.02	(0.05)	-0.17 *	(0.08)
Proportion living in poverty	0.25 **	(0.08)	0.05	(0.14)	0.17 *	(0.07)	-0.02	(0.12)
Proportion white	0.11	(0.21)	0.42	(0.44)	0.01	(0.19)	0.17	(0.36)
Proportion black	-0.12 *	(0.06)	-0.22 *	(0.09)	-0.09 +	(0.05)	-0.22 *	(0.09)
Proportion Asian	0.09	(0.11)	0.15	(0.17)	0.24 **	(0.09)	0.33 *	(0.17)

Online Table 9, Part 4 of 4: Propensity Score Estimates of Effect of Switching from Low- to High-Performing School on Teachers' Value-Added Scores

	Change in Value-Added Score[a]				Final Value-Added Score[b]			
	Teachers Unmatched		Teachers Matched		Teachers Unmatched		Teachers Matched	
	Models Adjusted by Baseline Covariates, continued							
	Model A2		Model A5		Model B2		Model A5	
	B	(SE)	B	(SE)	B	(SE)	B	(SE)
Characteristics of Students in Teacher's School, continued								
Proportion in Gifted and Talented	-0.15	(0.11)	0.00	(0.17)	-0.21 *	(0.09)	-0.09	(0.13)
Proportion in special education	-0.21 *	(0.08)	-0.03	(0.17)	-0.06	(0.07)	-0.07	(0.15)
Proportion in English Learner program	-0.04	(0.07)	-0.24 *	(0.10)	0.04	(0.06)	-0.18 *	(0.08)
School Resources								
Per pupil expenditures	0.15 +	(0.08)	-0.13	(0.23)	0.14 *	(0.07)	0.14	(0.21)
Average salary for teachers with 20 years or more ex	-0.04	(0.05)	0.00	(0.05)	-0.05	(0.04)	-0.03	(0.05)
Proportion core teachers with Master's or higher	-0.11 **	(0.04)	-0.04	(0.07)	-0.06	(0.04)	0.01	(0.05)
Average years of experience across all teachers	-0.39 ***	(0.08)	-0.37 *	(0.15)	-0.23 **	(0.07)	-0.24 *	(0.12)
Average years of experience across core teachers	0.32 ***	(0.08)	0.22	(0.14)	0.20 **	(0.07)	0.14	(0.12)
School is a magnet	0.08	(0.06)	0.12	(0.10)	0.03	(0.05)	0.09	(0.09)
Constant	-0.09	(0.11)	-0.19	(0.20)	-0.22 *	(0.10)	-0.18	(0.18)

Note: Analyses based on teachers who started in low-performing school. Estimates are from multilevel linear regression models with random-intercept at level of teacher's origin school.

a-These categories are not mutually exclusive.

b-The difference in teachers' scores between their last and first years in the district between 2007-08 and 2012-13, to assess the robustness of results in Figure 1.

c-Value-added score for last year in district between 2007-08 and 2012-13. Results adjusted by initial-value-added-scores and teachers' years of experience first year in district.

Online Table 10, Part 1 of 4: Propensity Score Estimates of Effect of Switching from High- to Low-Performing School on Teachers' Value-Added Scores

	Change in Value-Added Score[a]				Final Value-Added Score[b]			
	Teachers Unmatched		Teachers Matched		Teachers Unmatched		Teachers Matched	
	B	(SE)	B	(SE)	B	(SE)	B	(SE)
	Unadjusted Models							
Characteristics of Teachers	Model A1		Model A3		Model B1		Model B3	
Switched from high- to low-performing school	-0.27 ***	(0.06)	-0.29 **	(0.10)	-0.23 ***	(0.06)	-0.19 *	(0.09)
Constant	-0.02	(0.03)	0.05	(0.06)	0.05	(0.03)	0.04	(0.07)
	Models Adjusted by Propensity Scores							
			Model A4				Model B4	
Switched from high- to low-performing school			-0.29 **	(0.10)			-0.19 *	(0.09)
Propensity score			-0.05	(0.03)			0.00	(0.03)
Constant			0.00	(0.07)			0.04	(0.08)
	Models Adjusted by Baseline Covariates							
	Model A2		Model A5		Model B2		Model A5	
Switched from high- to low-performing school	-0.17 **	(0.06)	-0.21 *	(0.10)	-0.17 **	(0.06)	-0.23 *	(0.09)
Mean value-added score relative to school mean	-0.65 ***	(0.02)	-0.62 ***	(0.06)	0.35 ***	(0.02)	0.36 ***	(0.06)
Years of experience:								
3 or fewer (ref)	-		-		-		-	
More than 3, up to 8	-0.13 **	(0.05)	0.01	(0.09)	-0.12 *	(0.05)	0.03	(0.09)
More than 8, up to 15	-0.02	(0.06)	0.01	(0.09)	0.00	(0.06)	0.04	(0.09)
More than 15	-0.23 ***	(0.06)	-0.18	(0.11)	-0.22 ***	(0.06)	-0.16	(0.12)
Degree:[a]								
Math	0.23	(0.21)	-0.36	(0.33)	0.19	(0.20)	-0.80 *	(0.38)
Reading/Language arts	0.01	(0.10)	0.01	(0.31)	-0.01	(0.10)	0.01	(0.34)
Science	-0.03	(0.08)	0.08	(0.14)	-0.05	(0.07)	0.05	(0.14)
Social studies	0.03	(0.06)	0.08	(0.09)	0.03	(0.06)	0.12	(0.09)
Education	-0.13 *	(0.05)	0.05	(0.09)	-0.16 **	(0.05)	0.01	(0.08)

Online Table 10, Part 2 of 4: Propensity Score Estimates of Effect of Switching from High- to Low-Performing School on Teachers' Value-Added Scores

	Change in Value-Added Score[a]				Final Value-Added Score[b]			
	Teachers Unmatched		Teachers Matched		Teachers Unmatched		Teachers Matched	
	Models Adjusted by Baseline Covariates, continued							
	Model A2		Model A5		Model B2		Model A5	
Characteristics of Teachers, continued	B	(SE)	B	(SE)	B	(SE)	B	(SE)
Master's degree or higher	0.08 +	(0.04)	0.14	(0.10)	0.06	(0.04)	0.13	(0.10)
Certification focus:[a]								
Math	0.05	(0.07)	0.03	(0.17)	0.05	(0.07)	0.02	(0.15)
Reading/Language arts	-0.07	(0.05)	-0.07	(0.12)	-0.05	(0.05)	-0.04	(0.12)
Science	-0.24 **	(0.08)	-0.07	(0.23)	-0.21 **	(0.08)	-0.05	(0.22)
Social studies	-0.05	(0.07)	0.01	(0.18)	-0.05	(0.07)	0.06	(0.17)
Special education	-0.06	(0.09)	-0.27	(0.19)	-0.03	(0.08)	-0.25	(0.18)
English proficiency	0.01	(0.05)	-0.07	(0.08)	-0.01	(0.04)	-0.07	(0.07)
Male	-0.04	(0.05)	-0.03	(0.11)	-0.05	(0.05)	-0.08	(0.10)
Race:								
White, non-Hispanic (ref)	-		-		-		-	
Black, non-Hispanic	-0.13 *	(0.05)	-0.16	(0.10)	-0.11 *	(0.05)	-0.05	(0.11)
Hispanic, white	-0.08	(0.06)	-0.16	(0.12)	-0.05	(0.06)	-0.06	(0.11)
Asian, non-Hispanic	0.18 +	(0.09)	-0.38 *	(0.15)	0.14	(0.09)	-0.18	(0.15)
Other	-0.11	(0.10)	-0.10	(0.10)	-0.15 +	(0.09)	-0.04	(0.10)
Characteristics of Teacher's Students								
Number of students	0.05 +	(0.03)	0.11 +	(0.06)	0.04	(0.03)	0.07	(0.06)
Proportion male	0.00	(0.02)	0.05	(0.03)	0.01	(0.02)	0.05	(0.03)
Proportion white	-0.02	(0.04)	0.04	(0.12)	-0.02	(0.04)	0.06	(0.11)
Proportion Asian	0.02	(0.03)	0.04	(0.06)	0.02	(0.03)	0.02	(0.06)
Proportion other race	0.00	(0.02)	0.06	(0.04)	0.00	(0.01)	0.05	(0.04)

Online Table 10, Part 3 of 4: Propensity Score Estimates of Effect of Switching from High- to Low-Performing School on Teachers' Value-Added Scores

	Change in Value-Added Score[a]				Final Value-Added Score[b]			
	Teachers Unmatched		Teachers Matched		Teachers Unmatched		Teachers Matched	
	Models Adjusted by Baseline Covariates, continued							
	Model A2		Model A5		Model B2		Model A5	
Characteristics of Teacher's Students, continued	B	(SE)	B	(SE)	B	(SE)	B	(SE)
Proportion not economically disadvantaged	0.02	(0.03)	-0.10	(0.15)	0.01	(0.03)	-0.03	(0.12)
Proportion eligible for reduced lunch	-0.02	(0.02)	-0.03	(0.04)	-0.02	(0.02)	-0.04	(0.04)
Proportion eligible for free lunch	0.05	(0.04)	-0.03	(0.08)	0.01	(0.04)	-0.05	(0.07)
Proportion in English Learner program	-0.04	(0.04)	-0.02	(0.06)	-0.02	(0.04)	0.00	(0.05)
Proportion in Gifted and Talented	0.01	(0.02)	-0.04	(0.05)	0.01	(0.02)	-0.06	(0.05)
Proportion in special education	0.03	(0.02)	0.03	(0.05)	0.01	(0.02)	0.01	(0.05)
Proportion taking Spanish version of test	-0.04	(0.03)	0.00	(0.04)	-0.04	(0.03)	-0.01	(0.04)
Proportion tested for grade 3	-0.19	(0.31)	0.19	(0.79)	-0.45	(0.30)	-0.01	(0.68)
Proportion tested for grade 4	-0.17	(0.31)	0.19	(0.79)	-0.42	(0.30)	0.01	(0.68)
Proportion tested for grade 5	-0.17	(0.30)	0.23	(0.75)	-0.41	(0.29)	0.04	(0.64)
Proportion tested for grade 6	-0.19	(0.26)	0.10	(0.64)	-0.35	(0.25)	0.01	(0.55)
Proportion tested for grade 7	-0.10	(0.24)	0.15	(0.60)	-0.27	(0.23)	0.06	(0.51)
Proportion tested for grade 8	-0.11	(0.24)	0.14	(0.60)	-0.26	(0.24)	0.07	(0.51)
Characteristics of Students in Teacher's School								
Number of students	0.04	(0.05)	-0.04	(0.10)	0.01	(0.05)	-0.08	(0.11)
Proportion not economically disadvantaged	0.12	(0.13)	0.10	(0.28)	0.13	(0.13)	-0.09	(0.25)
Proportion eligible for reduced lunch	0.02	(0.04)	0.01	(0.09)	0.00	(0.04)	-0.02	(0.07)
Proportion living in poverty	0.00	(0.09)	0.02	(0.13)	0.03	(0.09)	-0.03	(0.13)
Proportion white	-0.05	(0.09)	-0.02	(0.24)	-0.08	(0.09)	0.00	(0.19)
Proportion black	-0.06	(0.06)	-0.14	(0.11)	-0.13 *	(0.06)	-0.12	(0.10)
Proportion Asian	-0.03	(0.04)	-0.13	(0.14)	-0.01	(0.04)	-0.10	(0.15)

Online Table 10, Part 4 of 4: Propensity Score Estimates of Effect of Switching from High- to Low-Performing School on Teachers' Value-Added Scores

	Change in Value-Added Score[a]				Final Value-Added Score[b]			
	Teachers Unmatched		Teachers Matched		Teachers Unmatched		Teachers Matched	
	Models Adjusted by Baseline Covariates, continued							
	Model A2		Model A5		Model B2		Model A5	
	B	(SE)	B	(SE)	B	(SE)	B	(SE)
Characteristics of Students in Teacher's School, continued								
Proportion in Gifted and Talented	0.01	(0.04)	0.13	(0.11)	0.02	(0.04)	0.12	(0.11)
Proportion in special education	0.07	(0.09)	-0.19	(0.24)	0.02	(0.09)	-0.19	(0.19)
Proportion in English Learner program	0.08	(0.06)	-0.01	(0.11)	0.03	(0.06)	-0.05	(0.09)
School Resources								
Per pupil expenditures	-0.05	(0.07)	0.08	(0.24)	0.07	(0.07)	0.03	(0.18)
Average salary for teachers with 20 years or more exp.	0.01	(0.03)	0.06	(0.06)	0.03	(0.03)	0.02	(0.05)
Proportion core teachers with Master's or higher	-0.08 *	(0.04)	-0.03	(0.06)	-0.06 +	(0.04)	-0.01	(0.05)
Average years of experience across all teachers	-0.04	(0.08)	0.22	(0.17)	-0.05	(0.08)	0.11	(0.16)
Average years of experience across core teachers	0.08	(0.08)	-0.16	(0.16)	0.05	(0.08)	-0.12	(0.15)
School is a magnet	0.13 *	(0.07)	-0.01	(0.11)	0.13 +	(0.07)	0.04	(0.10)
Constant	-0.02	(0.08)	0.11	(0.14)	0.16 *	(0.08)	0.15	(0.14)

Note: Analyses based on teachers who started in high-performing school. Estimates are from multilevel linear regression models with random-intercept at level of teacher's origin school.

a-These categories are not mutually exclusive.

b-The difference in teachers' scores between their last and first years in the district between 2007-08 and 2012-13, to assess the robustness of results in Figure 1.

c-Value-added score for last year in district between 2007-08 and 2012-13. Results adjusted by initial-value-added-scores and teachers' years of experience first year in district.