NATIONAL
CENTER for ANALYSIS of LONGITUDINAL DATA in EDUCATION RESEARCH

TRACKING EVERY STUDENT'S LEARNING EVERY YEAR

A program of research by the American Institutes for Research with Duke University, Northwestern University, Stanford University, University of Missouri-Columbia, University of Texas at Dallas, and University of Washington

AIR

Taking their First Steps: The Distribution of New Teachers into School and Classroom Contexts and Implications for Teacher Effectiveness and Growth

Paul Bruno

Sarah Rabovsky

Katharine Strunk

WORKING PAPER No. 212-0119-1 • January 2019

Taking their First Steps: The Distribution of New Teachers into School and Classroom Contexts and Implications for Teacher Effectiveness and Growth

Paul Bruno
University of Southern California

Sarah Rabovsky
University of Southern California

Katharine Strunk
Michigan State University/CALDER

Contents

Acknowledgments

We are grateful to the Los Angeles Unified School District for providing the data necessary to conduct this research, and in particular to Vivian Ekchian, Sergio Franco, Cynthia Lim and Emily Mohr for their partnership in developing the research agenda and Inocencia Cordova, Joshua Klarin, Jonathan Lesser and Kevon Tucker-Seeley for their assistance in obtaining and understanding the data.

This research was supported in part by the National Center for the Analysis of Longitudinal Data in Education Research (CALDER), which is funded by a consortium of foundations. For more information about CALDER funders, see www.caldercenter.org/about-calder. All opinions expressed in this paper are those of the authors and do not necessarily reflect the views of our funders, the Los Angeles Unified School District or the institutions to which the authors are affiliated.

CALDER • American Institutes for Research
1000 Thomas Jefferson Street NW, Washington, DC 20007
202-403-5796 • www.caldercenter.org

Taking their First Steps: The Distribution of New Teachers into School and Classroom Contexts and Implications for Teacher Effectiveness and Growth

Paul Bruno, Sarah Rabovsky, Katharine Strunk
CALDER Working Paper No. 212-0119-1
January 2019

Abstract

Novice teachers' professional contexts may have important implications for their effectiveness, development, and retention. However, descriptions of these contexts suffer from data limitations, resulting in unidimensional or vague characterizations. Using 10 years of administrative data from the Los Angeles Unified School District, we describe patterns of new teacher sorting using 27 context measures organized along three distinct dimensions - intensity of instructional responsibilities, homophily, and colleague qualifications – and use school-level survey data to measure a fourth dimension (professional culture). Relative to more experienced teachers, novice teachers have placements that are more challenging along the first three dimensions, and composite measures are differentially predictive of teachers' outcomes. This suggests that policymakers should consider placements to better retain and develop novice teachers.

Keywords: Teacher labor markets, teacher sorting, early career teachers, school context

Introduction

In recent years, conversations about teachers have largely focused on teacher recruitment (e.g., Maranto & Shuls, 2012; Will, 2017), identifying effective teachers (e.g., Burnette, 2017; Hanushek & Rivkin, 2012), and holding teachers accountable for outcomes (e.g., Burnette, 2017; Dee & Wyckoff, 2015). While attention to these concerns stems directly from evidence about the importance of high quality teachers for students' short- and longer-term outcomes (e.g., Chetty, Friedman, & Rockoff, 2014; Gershenson, Holt, & Papageorge, 2016; Kraft, 2017) and concerns about teacher shortages (e.g., Dee & Goldhaber, 2017; Podolsky, Kini, Bishop, & Darling-Hammond, 2016), these discussions often gloss over considerations of teacher retention and development. Retention and support of teachers, especially novice teachers, has critical implications for school operations and, in turn, student learning and achievement. For instance, teacher turnover imposes strains on school operations (Ronfeldt, Loeb, & Wyckoff, 2013) and district resources (Milanowski & Odden, 2007). Because early-career teachers are substantially more likely to exit their schools and districts than are their more experienced colleagues (Keigher, 2010), efforts to improve novice teachers' working conditions and support them through their first years of service are particularly important.

A growing literature documents that novice teachers often experience more difficult initial working conditions and school contexts than do their more senior colleagues (Feng, 2010; Goldhaber, Lavery, & Theobald, 2015; Kraft & Papay, 2014), and that these contexts are likely associated with early-career teachers' success and persistence in their jobs and the profession (Feng, 2010; Kraft & Papay, 2014). However, for the most part, efforts to characterize these contexts and isolate their most salient features suffer from limitations in the administrative datasets used and from potential biases stemming from reliance on teachers' self-reports.

In this study, we provide detailed evidence about the working conditions faced by new teachers in their classrooms and schools and relate these classroom and school characteristics to teachers' professional trajectories. We begin by offering a conceptual framework within which teachers' placements can be characterized along four dimensions with plausibly distinct implications for teachers. Using both school-level survey data and administrative data linking teachers and administrators to students and schools, we then characterize teachers' contexts during a 10-year period in the Los Angeles Unified School District along these four dimensions: whether teachers' placements (1) impose greater demands on teachers as a consequence of greater *instructional load*; (2) allow for greater *homophily* due to social congruence between teachers and their students, fellow teachers, and school leaders; (3) provide opportunities and supports to teachers due to *colleagues with stronger qualifications*; and (4) have stronger or weaker *professional culture*. We show that these four dimensions are largely distinct from one another on average, and then consider two research questions related to these measures of context: (1) *How do the professional contexts of novice teachers differ from those of more experienced teachers in the same district and schools?;* and (2) A*re the contexts in which teachers work associated with their effectiveness, development, and retention?*

We find that relative to veterans, novice teachers, and especially teachers in their first two years in the classroom, are placed in classrooms and schools with higher instructional loads, less homophily, and lesser-qualified colleagues, though we find few differences in their professional cultures as measured by aggregate survey responses. Further, we find that, on average, teachers in placements with higher levels of instructional load have smaller contributions to student achievement on standardized tests, lower evaluation ratings, and lower attendance. Improvements in all four of our context measures uniquely predict higher levels of teacher

retention, due primarily to lower levels of school switching for veterans and lower rates of district exit among novices. Importantly, as with most of the extant research on novice teacher development, because they are not identified from exogenous variation in teachers' contexts these results do not have a clear causal interpretation. Nevertheless, results point to avenues for future work and potential levers for change available to policymakers and school administrators.

In the remainder of this paper we begin with a discussion of the existing literature on the sorting of new teachers into schools and classrooms and the importance of those placements for teacher effectiveness and retention.[1] Next, we outline our data and empirical strategies and then present the results of our analyses. We conclude with a discussion of the implications for policymakers and for future research.

Framework and Background

The research base documenting relationships between teachers' experience levels and their teaching contexts includes numerous studies that consider very different characteristics of those contexts. However, this literature is often limited in at least three ways. First, sample size and data limitations rarely allow individual studies to consider more than a few aspects of teachers' contexts at a time. Second, and perhaps relatedly, while the aspects of teachers' contexts that can be plausibly measured and analyzed has grown in recent years, this has generally not been accompanied by an effort to relate these measures to one another. This limits the practical and theoretical use of many results because it is not always clear whether different context measures operate through similar mechanisms, have different consequences, or should be expected to correlate with one another. Third, when multiple aspects of teachers' working

[1] We use terms like "sorting" and "placement" interchangeably when referring to the processes by which teachers arrive in schools and classrooms. In the analyses that follow, as in much of the prior literature discussed below, we are not able to disentangle teacher preferences from those of administrators or from formal rules (e.g., in collective bargaining agreements) that determine teacher assignment.

conditions have been simultaneously considered, this has generally been done using survey-based measures, which may suffer from reporting biases and often lack clear links to specific aspects of teachers' jobs that can be directly observed or manipulated by administrators. We therefore begin by articulating a conceptual framework in which teachers' placements can vary along four dimensions that theory or evidence indicate may matter for teachers. This framework then helps to organize our review of the literature and to motivate our subsequent analyses, which incorporate both survey-based measures of teachers' working conditions and a wide range of more readily-observable characteristics of teachers' students and coworkers using standard administrative data.

Conceptual Framework

Among the most commonly-studied aspects of new teachers' contexts are the characteristics of their students. These findings are frequently framed as informing our understanding of the opportunities enjoyed by students by virtue of the experience levels, and thus the effectiveness, of their instructors (e.g. Goldhaber et al., 2015). Additionally, these contextual features are often assumed to have important implications for teachers themselves because they indicate placements that are challenging (e.g., because students require more active classroom management). The first dimension of our framework thus is comprised of features of teachers' contexts that indicate aspects of *instructional load* that may impose greater professional demands on teachers, and which may therefore make a position less attractive or less conducive to effective teacher practice.

Research on teachers' placements often assumes that the salient features of teaching placements are defined by characteristics of students (e.g., their prior achievement) and are thus largely independent of the particular teachers who occupy those placements. There are reasons to

believe, however, that salient features of teachers' environments will in some cases be determined by the interaction of teacher and school characteristics. In particular, sociologists have identified *homophily* – that is, attraction between parties based on similarity – as an important driver of social network formation, with similarity of race or gender appearing to be particularly significant (McPherson, Smith-Lovin, & Cook, 2001). Though not as well studied in the educational literature as aspects of instructional load, some evidence suggests that similarity between teachers and their schools may be important for the success of both (e.g., Egalite, Kisida, & Winters, 2015; Strunk & Robinson, 2006), and homophily thus represents the second dimension in our framework.

More recently, researchers have begun to examine the importance of the quality of teachers' coworkers and teacher collaboration for individual teachers' success. The motivation for such studies is straightforward: teachers who have higher-quality (e.g., harder-working or more highly-skilled) coworkers may receive more support in facing the demands of their work or be held to higher standards, and large majorities of teachers report that colleagues are important to their own success (MetLife Foundation, 2009). We thus define *colleague qualifications* as a third dimension of teachers' contexts, including measures that may indicate the quality of the peers with which teachers work and on whom they may be able to draw. Yet we also distinguish the qualifications of teachers' coworkers from the quality of professional interactions between staff, with the latter constituting *professional culture*, our fourth dimension.

This framework offers two advantages for present purposes. First, as discussed below, it guides our analyses, including our selection of measures of teachers' contexts. Second, it allows us to impose order on previous research, including relating disparate studies to one another and identifying dimensions of teachers' working conditions that are relatively under-studied. Our

framework thus organizes our review of the literature, and within each dimension we consider not only findings relating teachers' experience levels to their contexts, but also, when possible, research linking those contexts to teachers' outcomes.

Previous Research

Instructional load. Compared to those of more experienced peers, placements for new teachers are characterized by higher proportions of low-income, minority, and low-achieving students (Clotfelter, Ladd, & Vigdor, 2005; Goldhaber et al., 2015; Kalogrides, Loeb, & Béteille, 2013; Lankford, Loeb, & Wyckoff, 2002; Scafidi, Sjoquist, & Stinebrickner, 2007), disparities that are even more pronounced in urban areas (Lankford et al., 2002). This type of sorting occurs both between and within schools; compared to their colleagues at the same school, new teachers are more likely to teach in classrooms with greater proportions of minority, low-income, and lower-achieving students and students who require special education or English language learning services (Clotfelter et al., 2005; Feng, 2010; Goldhaber et al., 2015; Rothstein, 2009). New teachers are also placed in classrooms with larger proportions of students with disciplinary infractions (Feng, 2010) and new high school teachers are more often assigned to teach higher numbers of ninth grade students (Neild & Farley-Ripple, 2008) and lower-level courses (Kelly, 2004).

Contexts and outcomes. More difficult placements of the sort described above are associated with higher levels of teacher attrition. New teachers in particular are more likely to exit if placed with lower achieving students with more discipline problems (Donaldson & Johnson, 2010; Feng, 2010; Kukla-Acevedo, 2009). Teachers' placements may also impose demands that are unrelated to student characteristics *per se*. For example, there is evidence that in addition to becoming more effective with experience in general, teachers also acquire skills

specifically related to the curricula with which they have experience (Ost, 2014). Thus, the courses to which they are assigned may have implications for teachers' effectiveness over and above the characteristics of the students enrolled in those courses.

Homophily. We are not aware of prior research directly examining whether teachers of different experience levels enjoy different degrees of homophily in their teaching placements, but some studies suggest that this may be the case. For example, Boyd, Lankford, Loeb, Ronfeldt, and Wyckoff (2011) find that black and Hispanic teachers in New York City are less likely to apply for transfers out of schools with larger shares of same-race students. Moreover, there is evidence that when teachers transfer, they transfer into schools with larger shares of same-race students (e.g., Hanushek & Rivkin, 2007). These results may imply that racial congruence between teachers and their schools will tend to be higher among teachers with higher experience levels, either because more experienced teachers have had more opportunities to sort into homophilic environments or because more experienced teachers may enjoy greater transfer rights under local collective bargaining agreements by virtue of their accumulated seniority (e.g., Goldhaber, Lavery, & Theobald, 2016).

Context and outcomes. Unlike the case of instructional load, where teacher sorting patterns related to experience are generally viewed as a source of concern, a small but growing body of research suggests that the sorting of new teachers of color to placements with higher proportions of students of color may have advantages. For example, research on teacher-student race congruence finds that teachers have higher expectations for their same-race students (Gershenson et al., 2016) and evaluate those students' behaviors more favorably (Downey & Pribesh, 2004; Wright, Gottfried, & Le, 2017), and that students make larger achievement gains when assigned to same-race teachers (Dee, 2004; Egalite, Kisida, & Winters, 2015). In addition to apparent

benefits for students, there is evidence that teachers are more likely to remain in their school placements when student racial compositions more closely match their own (Hanushek, Kain, & Rivkin, 2004; Strunk & Robinson, 2006). Racial congruence between teachers and other school staff may also be important; having a same-race administrator may increase job satisfaction and reduce turnover among teachers (Grissom & Keiser, 2011).

Colleague qualifications and professional culture. The extent to which teachers of different experience levels are surrounded by colleagues of different quality is not well documented. Lower-performing schools, where many new teachers are placed, generally have less-experienced principals and principals who are rated less effective according to teacher surveys (Grissom, 2011). High rates of teacher turnover are more common in lower-performing schools as well, limiting the amount of peer expertise to which new teachers are exposed (Boyd, Lankford, Loeb, & Wyckoff, 2005).

However, the observed quality of teachers' colleagues may be distinct from the frequency and quality of staff interaction and collaboration. Survey evidence indicates that teachers of different experience levels are no more or less likely to report that their schools engage in higher or lower levels of collaboration (MetLife Foundation, 2009), though less (more) experienced teachers report that the collaboration in which they engage is higher- (lower-) quality in at least some cases (Ronfeldt, Farmer, McQueen, & Grissom, 2015). Given the self-reported nature of these data, it is difficult to know whether these differences reflect genuine variation in collaboration quality, let alone true differences in colleague quality, and in the analyses below we treat observable staff characteristics and survey measures of professional behavior as reflecting distinct features of teachers' professional environments.

Context and outcomes. There is suggestive evidence that the other staff at a school matter for schools' professional communities and new teachers' trajectories. For example, higher-quality collaboration is associated with more rapid teacher improvement and larger student achievement gains (Ronfeldt et al., 2015), pointing to the potential value of higher-quality peers with whom to collaborate. Kraft & Papay (2014) find that a composite measure of school professional environment, including such teacher-reported factors as peer collaboration, evaluation quality, and school discipline, predicts faster teacher improvement. Similarly, teacher ratings of their administrators and relationships among colleagues are significant predictors of teacher satisfaction and retention as well as student achievement growth (Boyd, et al., 2011; Grissom, 2011; Johnson, Kraft, & Papay, 2012; Kraft, Marinell, & Yee, 2016). These studies, however, do not directly link teachers' outcomes to specific attributes of their coworkers.

A few studies suggest that school-level professional environments may be at least partially explicable in terms of the observable characteristics of the staff. New teachers in particular make larger contributions to student achievement when working with colleagues who have made larger contributions in the past (Jackson & Bruegmann, 2009). Additionally, as the percent of teachers in a school with more experience, advanced degrees, or professional certification grows, the likelihood that a teacher moves within the district falls (Feng & Sass, 2017). Furthermore, when teachers enter a new school their attendance rates converge to those of the school (Ost & Schiman, 2017), suggesting that coworkers shape professional norms around effort in addition to whatever effects they may have on absolute performance or growth.

Summary

Existing research indicates both that teachers with different levels of experience tend to have different professional contexts on average, and that these differences are likely to have

implications for new teachers' outcomes. However, as discussed above, this literature is often limited, considering few aspects of teachers' placements at a time, relying heavily on survey-based measures of working conditions, and lacking criteria by which different dimensions of teachers' working conditions can be defined or measured. In addition to offering a novel and practical conceptual framework, described above, we address these limitations by exploiting a longitudinal dataset from the second-largest school district in the country, allowing for the simultaneous measurement of many features of teachers' professional contexts.

Data

The data for this project largely come from administrative records provided by the Los Angeles Unified School District (LAUSD). These records primarily span the 2007-8 through 2016-17 school years and link administrators and teachers to individual schools and students. We additionally use student-level data from 2006-7 for student prior-year data in 2007-8 and teacher data from 2017-18 to code teacher mobility after the 2016-17 school year.

School Data

Effectively tracking patterns of teacher sorting within a district requires both a complete survey of schools and consistent data reporting across those schools. We thus include in our sample all schools in LAUSD with a few exceptions. We exclude early childhood and adult education centers, which typically do not employ certificated K-12 teachers. We also exclude independent charter schools, which are largely autonomous and because LAUSD does not have authority to share data from independent charter schools with external entities. However, other charter schools with tighter operational links with the district are included (labeled "affiliated" charter schools), as are traditional public schools offering alternative K-12 instructional environments (e.g., continuation schools and schools for pregnant minors).

We merge these data with school-level survey data from the annual School Experience Survey (SES) administered by LAUSD since the 2012-13 school year.[2] The survey is available to all staff assigned on at least a half-time basis to a school site. The district typically releases responses, aggregated to the school level, publicly online if at least 11 staff respond. To ensure that responses accurately reflect the views of school staff we further exclude data on questions to which less than 50 percent of surveyed staff at a site responded. Ultimately, we include 808 unique schools in our sample, the characteristics of which are summarized in Table 1.

Administrator Data

District administrative data files link administrators, including both principals and assistant principals, with individual school sites as of "norm day," typically the fifth Friday of the school year, on which student enrollments are documented for resource allocation purposes. These files identify each administrator's race, gender, years of experience as an administrator, and highest degree held. Because administrative responsibilities, such as teacher evaluation, are often shared between principals and assistant principals, we retain both and treat them similarly (e.g., gender congruence between a teacher and a principal is treated in the same way as a similar congruence between a teacher and an assistant principal), though results are very similar if we exclude assistant principals. Table 1 summarizes school-level administrator characteristics. The "average" school has one principal and in a few cases no principal is reported at all, though in these cases there is always at least one assistant principal assigned to the school.

Student Data

District records link students to schools, also as of norm day in each year. These files include student race and gender and indicators of whether students are English language learners

[2] The SES has been administered in some form since the 2008-9 school year, but the specific questions we use here have been asked on the survey consistently only since 2012-13.

or are eligible for free or reduced-price lunch or special education services. Student data files also include students' academic and non-academic outcomes. Academic outcomes include students' performance on statewide standardized tests in math and ELA and, in some years, science and social studies, which we standardize within test and year. We also observe students' course and grade records, from which we calculate grade point averages (GPAs), both overall and in specific subjects: math, ELA, social studies, science, physical education, and art. GPAs are standardized within grade, year, and, for subject-specific GPAs, subject area. Non-academic outcomes include student attendance rates and suspension records.

Student characteristics are summarized at the school level in Table 1. The schools in our sample enroll an average of 755 students each year, but there is considerable variation; the smallest schools – typically continuation schools – enroll as few as 29 students and the largest – comprehensive senior high schools – enroll more than 4,000 in at least some years. These schools' student bodies vary widely as well. For example, the "average" student body at these schools is 73 percent Hispanic and 76 percent eligible for free- or reduced-price lunch, but we observe some schools with nearly no such students.

Teacher Data

District norm day reports provide gender, race, and job title information on all staff, and link staff to schools. It is in this way that we associate teachers with individual school sites each year and across time. We retain in our data all elementary, secondary, or special education teachers unless they hold a part-time or temporary position (e.g., as substitutes). For teachers, the district also reports years of teaching experience (truncated at 10 for teachers in their tenth or later year of teaching) and the highest degree held. We observe 40,879 unique teachers over this

period, with their characteristics described in Table 2. Just over half are elementary teachers, more than two-thirds are female, and 40 percent are white.

Teacher outcomes. We consider four types of outcomes for teachers. First, beginning in 2012-13, we have measures of total hours absent. Absences are distinguished based on whether they were for legally protected reasons (e.g., military leave or jury duty) or unprotected reasons (e.g., personal days or ordinary sick leave). We focus on unprotected absences, which are more discretionary from a teacher's perspective, and use as our outcome each teacher's attendance rate, or the share of hours a teacher was scheduled to work (excluding protected absences) during which she attended work. Teachers attend approximately 96 percent of their scheduled work hours on average by this measure, or 173 days in a 180-day school year.

Second, student-level achievement data allows for estimates of math and ELA teachers' value-added measures (VAMs) of their contributions to student achievement. We estimate:

$$(1) \qquad Ach_{ijst} = \beta_1 Ach^{math}_{ijst-1} + \beta_2 Ach^{ela}_{ijst-1} + \boldsymbol{X}_{ijst}\boldsymbol{\theta} + \boldsymbol{T}_{jt}\boldsymbol{\Omega} + \varepsilon_{ijst}$$

where *Ach* is either math or ELA achievement for student *i* with teacher *j* in school *s* in year *t*, and we control for achievement in the prior year in both subjects. $\boldsymbol{X}$ is a vector of student characteristics, including indicators of student race, gender, free- or reduced-price lunch eligibility, special education status, English learner status, and grade level. Teachers' VAMs are estimated by the coefficients on a set of teacher fixed effects ($\boldsymbol{T}$). ε is an error term, and we estimate VAMs separately for elementary and secondary students and for each year. Students can be linked to multiple teachers in a given year, with each student-teacher link weighted by the fraction of the year for which the link was observed. We estimate heteroskedasticity-robust standard errors and use standard empirical Bayes shrinkage methods to account for uncertainty for teachers with few students or students with difficult-to-predict achievement trajectories

(Koedel, Mihaly, & Rockoff, 2015). VAMs are then standardized across all teachers in the district in each year. For additional information about VAM estimation, see Online Appendix B.

Third, we incorporate subjective performance evaluations of individual teachers, provided by LAUSD's Human Resources Division. The use of these evaluation data is complicated by the fact that the district transitioned from the previous teacher evaluation system – the Stull Evaluation Process – to the current Educator Development and Support: Teachers (EDST) system in 2013-14. Under both systems teachers are given a final overall rating indicating whether they "meet standards" or are "below standard" based primarily on classroom observations, with probationary teachers evaluated annually and tenured teachers evaluated every other year or, with sufficient satisfactory experience, as infrequently as every five years. However, the Stull and EDST systems differ in the standards against which teachers are evaluated and beginning in 2015-16 teachers could also earn a final rating of "exceeds standard". We use the receipt of a below standard rating as a teacher outcome that should be approximately comparable across teachers and years, and in the analyses below we also control for average districtwide changes over time to make within-year comparisons between teachers. Additionally, under the EDST system teachers are rated on a three-level (or four-level, in some years) scale on specific elements of their practice. We convert these ratings to numerical scores that we then average for each teacher and standardize across teachers in each year to produce a more continuous measure of teachers' observation outcomes from 2013-14 through 2016-17.

Finally, a common concern among both researchers and policymakers is the extent to which teachers can be retained to avoid costly and disruptive staff turnover. We therefore analyze whether teachers switch schools within the district from one year to the next or whether they leave the district altogether. We indicate teachers as switching schools at the end of the year

if on the next year in which they are observed they are working at a different school site, and as leaving the district if they are never subsequently observed in a district school. However, we do not indicate mobility outcomes for teachers who are laid off, as layoffs are conducted according to seniority criteria and are thus not discretionary for either teachers or administrators.

These teacher outcomes are summarized in Table 2, both for teachers in their first five years of teaching (whom we label "novice" in Table 2) and for those with six or more years of teaching experience (whom we call "veterans"). On average, relatively new teachers have higher attendance rates – missing approximately one less day of work per year than other teachers – and are less likely to be given a below-standard evaluation rating. On the other hand, new teachers have relatively low ELA VAMs and are more likely to switch schools or leave the district.

Characterizing Teacher Contexts

Instructional load. Guided by the literature and conceptual framework discussed above, we generate measures that plausibly proxy for possible academic, social, emotional, or behavioral needs of students that will tend to impose specific professional demands on the teachers to whom they are assigned. To minimize the possibility that students' apparent need is impacted by the teacher to whom they are assigned in a given year, we often define these measures in terms of students' characteristics in the prior year. These characteristics are summarized in the top panel of Table 3.

Because lower-achieving students are likely to require greater instructional supports, we include as measures of instructional load (1) students' prior year average performance averaged across all available tests and, when available for secondary teachers, (2) students' performance on the teacher's own subject area test. Similarly, greater diversity in prior achievement may impose greater instructional demands on teachers, for example because students require greater

instructional differentiation. We therefore include (3) the variation of students' prior achievement, defined as the mean of the standard deviations of their z-scores on their math and ELA tests and, for applicable secondary teachers and (4) the standard deviation of students' z-scores on the teacher's subject area test. Because standardized tests do not capture all relevant measures of students' academic performance or difficulties, we also include two GPA measures, including (5) students' average overall GPA in the prior year, standardized as described above and, for applicable secondary teachers, (6) students' average GPA in the teacher's own subject.

We also include two non-academic measures of students' performance. Because students who miss school more frequently may require additional supports or accommodations from teachers (Allensworth & Easton, 2007; Gottfried, 2014), we include (7) the average percent of school days students missed in the prior year. And because students with prior disciplinary infractions may be more likely to engage in disruptive misbehavior (e.g., Figlio, 2007), we include (8) the percent of students who were suspended at any point in the prior year. We also include three student characteristics that, while not direct measures of student performance, may be indicative of distinctive instructional need. These are the shares of students who are (9) eligible for free or reduced-price lunch, (10) eligible for special education services, or (11) classified as English learners. Finally, teachers may have different workloads because they are assigned more students or more unique classes to teach. We therefore include (12) average class size. For elementary teachers, class size is defined as the average number of students for which the teacher is responsible in each of the trimesters for which they are present at the school. For secondary teachers, class size is defined as the average number of students in each class period in each semester that they teach at the school. Additionally, for secondary teachers we include (13) the number of course preparations (i.e., the number of distinct courses that teachers are required

to teach) over the course of the year. Because teaching the second semester of a year-long course entails more preparation (i.e., to teach new content) than teaching a one semester course twice in one year, each semester of a year-long course is counted as a separate preparation.

Homophily. As discussed above, because teachers are heterogeneous, any placement, even if challenging in some objective sense, may be more attractive or supportive for some teachers than others due to aspects of congruence between the teacher and the school. Extant literature on relevant aspects of congruence in schools points to race and gender as plausible candidates. As racial homophily measures, we include (1) the share of students or (2) other site teachers who are the same race as the teacher, and (3) an indicator of whether the teacher has a same-race administrator (whether a principal or assistant principal). We also consider three analogous measures of gender congruence, for six homophily measures in total, summarized in the second panel of Table 3.

Colleague qualifications. Because more capable and committed coworkers may be more helpful for new teachers, we consider eight plausible proxies for those attributes in their administrators and fellow teachers. These are summarized in the third panel of Table 3. This includes the share of other site teachers who (1) are in their tenth or later year of teaching, (2) have a graduate degree, or (3) have National Board Certification. As more direct measures of teacher performance, we also include for each teacher (4) the share of other evaluated site teachers given a "below standard" final evaluation rating, (5) the mean attendance rate of other site teachers, and (6) the mean VAM of other site teachers, using the mean of ELA and math VAMs when a teacher has both, all measured in the prior year. We observe fewer characteristics of teachers' administrators, but include (7) the mean years of experience of site administrators

and, because virtually all administrators in LAUSD possess at least a master's degree, (8) an indicator of whether any site administrator has a doctorate.

Professional culture. We measure a school's "professional culture" by identifying 10 questions from the School Experience Survey that address the frequency and quality of professional interactions at teachers' school sites. These items are summarized in the bottom panel of Table 3. Because results are aggregated to the school level, we use as our professional culture measures the shares of site staff who agree or strongly agree with statements like, "This school promotes trust and collegiality among staff" or who say that "school leadership provide[s] useful feedback to you based on their observations" always or often. For a full list of survey questions, see Online Appendix Table A1.

Composite measures. Because our goal is to attempt to characterize new teachers' contexts based on levels of instructional load, homophily, colleague qualifications, and professional culture, we combine each set of the individual measures discussed above into a single composite measure for each of the four dimensions, provided in the last row of each panel in Table 3. Since the individual characteristics are constructed on different scales (e.g., z-scores vs. percentages), we follow a four-step process to combine them into a single construct. First, each individual characteristic is transformed to be measured in the same direction, such that more positive numbers indicate more of the dimension in question. Thus, for example, test scores and GPAs are multiplied by negative one such that lower true values correspond to greater instructional load. Second, each characteristic, no matter how initially measured, is standardized to have a mean of zero and standard deviation of one across all teachers in each year in the sample. Third, each teacher is given a composite score for each dimension in each year that is the mean of the standardized individual measures for that dimension. Despite the standardization of

each individual measure, these mean dimension scores do not retain a unit standard deviation and so to facilitate interpretation as a final step each composite measure is standardized again across teachers within each year.

Importantly, not every context measure can be constructed for every teacher. Teachers' composite dimension scores are thus constructed using different numbers of individual context measures, if they can be constructed at all. So, for instance, teachers' composite colleague qualification measures are estimated using as few as five or as many as eight colleague qualification measures. Because teachers' contexts are characterized using different quantities of information, in the models below when composite measures are used to predict teachers' outcomes, teacher observations are weighted in proportion to the number of individual component measures from which their composite scores are estimated, though in practice this makes little difference as most component measures are observed for most teachers.

These four composite scores are only weakly correlated with one another, with correlations ranging from $r = -.15$ (between the colleague qualifications and instructional load composites) to $r = .14$ (between the colleague qualifications and professional culture composites). Thus, to the extent that they usefully characterize dimensions of teachers' contexts, they also capture distinct information about those contexts. Though these weak relationships are perhaps surprising, this is also consistent with prior work finding that survey-based measures of three aspects of school culture – teacher autonomy, student behavior, and administrative support – were not highly correlated (Kukla-Acevedo, 2009).[3]

Methods

To compare the contexts of novices and veterans (RQ1), we estimate via OLS:

[3] Factor analytic techniques do not provide clean latent factors aligned to the content of our four context measures. We therefore use our index measures. Results from these analyses are available upon request.

$$(2) \qquad context_{ist} = \alpha_0 + \alpha_1 year\_1_{ist} + \alpha_2 year\_2_{ist} + \alpha_3 years\_3\_to\_5_{ist} + \gamma_t + \varepsilon_{ist}$$

Here *context* is a component or composite measure of the context, as described above, for teacher *i* in school *s* in year *t*. *year_1*, *year_2*, and *years_3_to_5* are dummy variables indicating teachers in their first, second, or third through fifth years of teaching, respectively, with coefficients estimating the difference in context between these groups and teachers in their sixth or later year of teaching. γ is a set of year dummies, and we cluster standard errors at the school-by-year level, since several of our context measures are defined at that level.

Our instructional load measures and some of our homophily measures are constructed using each teacher's students, rather than using other staff at the school site. Differences in these measures between teachers will reflect both differences arising because teachers work in different schools and differences arising because teachers within the same school have different assignments. For these student-based measures we therefore also estimate a variant of model (2) that includes a school-by-year fixed effect, comparing only teachers working within the same school in the same year, isolating differences between teachers of different experience levels that arise within schools.

Because homophily measures are defined by the interaction of teachers' characteristics with those of their placements, we estimate model (2) with and without controls for teachers' own race or gender when considering racial or gender homophily, respectively. For all four dimensions we estimate model (2) separately for elementary, secondary, and special education teachers, since teachers may sort differently when they hold different kinds of teaching positions.

Finally, to determine whether there are associations between our composite measures of teachers' contexts and their outcomes (RQ2), we estimate regressions predicting those outcomes,

interacting composite context scores with indicators of whether teachers are in their sixth or later year of teaching to allow for different relationships for teachers with different experience levels:[4]

$$(3)\ outcome_{ist} = \theta_0 + \theta_1 inst\_load_{ist} + \theta_2 inst\_load_{ist} * veteran_{ist} + \theta_3 homophily_{ist} + \theta_4 homophily_{ist} * veteran_{ist} + \theta_5 colleague\_qual_{ist} + \theta_6 colleague\_qual_{ist} * veteran_{ist} + \theta_7 prof\_culture_{ist} + \theta_8 prof\_culture_{ist} * veteran_{ist} + \boldsymbol{X}_{ist}\varphi + \gamma_t + \epsilon_{ist}$$

Here *outcome* is each (new) teacher's attendance rate, ELA or math VAM, average EDST rating, the probability of receiving a below standard evaluation rating, or the probability of leaving their school or the district. For continuous outcome measures we estimate our models using OLS, and we estimate categorical outcomes using logistic specifications, though results are very similar if linear probability models are used. *inst_load*, *homophily*, *colleague_qual*, and *prof_culture* are the composite measures of our context dimensions as defined above, and observations are weighted in proportion to the number of context measures used to construct the composites. $\theta_1, \theta_3, \theta_5,$ and θ_7 thus estimate relationships between composite context measures and outcomes for novices, with the coefficients on the interaction terms estimating the extent to which those relationships differ for veterans. Because the composite context measures are largely orthogonal to one another, results are very similar whether they are included individually or simultaneously in the model.

$\boldsymbol{X}$ is a vector of teacher characteristics, including indicators of teachers' gender, race, and years of experience. Additionally, we combine elementary, secondary, and special education teachers, and thus also include in $\boldsymbol{X}$ indicators for these job titles. Results are mostly similar for

[4] In results available upon request, we also estimate a variant of model (3) that interacts composite context scores with all three of the experience dummies used in model (2). We do not observe clear or consistent differences between the coefficients on each of those interaction terms, but those estimates are also much less precise, with standard errors that are in some cases more than twice as large as what is observed using a single dummy variable to indicate veterans.

these different kinds of teachers, and we present estimates for results for each group separately in Online Appendix Tables A3 and A4. γ is again a set of year dummies, and standard errors are again clustered at the school-by-year level.

As discussed above, an important consideration when estimating relationships between teachers' contexts and their outcomes is that teachers are not randomly assigned to working conditions. Estimates like those in model (3), and in much of the previous literature, may therefore be biased by unobserved differences between teachers if those differences are also related to their placement contexts. We exploit the longitudinal nature of our data to estimate variants of model (3) that include a teacher fixed effect, effectively comparing teachers to themselves in different years, as they experience different contexts over time. While this by no means guarantees that we isolate entirely exogenous variation in teachers' contexts, it eliminates fixed differences between teachers as a source of bias.

Results

RQ1: How do the professional contexts into which novice teachers sort differ from those of more experienced teachers in the same district and schools?

Below we discuss differences in instructional load, homophily, and colleague qualifications between teachers of different experience levels. We observe few differences in professional culture, as measured by survey questions, between teachers of different experience levels. Given space constraints, we present results related to professional culture in Online Appendix Table A2. Results for the other three dimensions are presented in Tables 4 through 6. Results for each type of teacher – elementary, secondary, or special education – are presented in separate panels in each table. In each subsection below, we first discuss comparisons of novice versus more experienced teachers throughout the district (reporting relationships from our

regressions with year fixed effects) and then discuss differences (if any) between these outcomes and those from models that instead include school-by-year fixed effects.

Instructional load. As shown in Table 4, novice teachers have considerably higher overall instructional loads than do their more experienced peers, with the largest gaps occurring between first year teachers and teachers with six or more years of experience. Even at the elementary level, where differences between novice and more experienced teachers are generally smallest, experience gaps on individual context measures accumulate to composite instructional load scores (column 1) that are 13 percent of a standard deviation larger for first- and second-year teachers than for veterans and the gap is twice as large for first year secondary teachers. These gaps shrink among more-experienced novices; across all three school levels, composite instructional load gaps with veterans are less than half as large for teachers in their third, fourth, or fifth years as for teachers in their first year.

For all three types of teacher – that is, for elementary, secondary, and special education teachers alike – composite instructional load gaps are driven at least in part by gaps in students' prior achievement and discipline records. Novice teachers have students with prior achievement that is between eight and 22 percent of a standard deviation lower on average than that of veteran teachers' students (column 2). Gaps are similar in magnitude when considering students' prior GPAs (column 6) or, for secondary teachers, prior achievement in each teachers' own subjects (columns 3 and 7). Similarly, novice teachers have slightly larger shares than veterans of students who were suspended in the previous year (column 11), by at least 0.37 percentage points for elementary teachers and as much as 1.9 percentage points for secondary teachers. Though not large in absolute terms, these gaps are large in proportional terms, since only one (five) percent

of the students of the average elementary (secondary or special education) teacher were suspended in the previous year.

Despite these gaps there are some cases in which differences in instructional load are negligible. For example, newer teachers do not have substantially larger class sizes than veterans (column 13), and novice secondary teachers have slightly smaller class sizes and fewer preparations (column 14). Nor do novice teachers have substantially larger shares of special education students than veterans. Thus, while our results align with previous work finding that the placements of novice teachers appear relatively more challenging than those of veterans, they also highlight that those placements are not necessarily more challenging in every respect.[5]

Comparing results in Table 4 with only a year fixed effect to those with a school-by-year fixed effect indicate that, consistent with the prior work discussed above (e.g., Goldhaber et al., 2015), both within- and between-school sorting play a role in producing the instructional load gaps that we observe. Composite instructional load gaps shrink by roughly half for all three types of teacher when novices are compared only to veterans in the same school in the same year, with similar patterns for several individual measures, including students' prior achievement and suspension rates.

Homophily. Table 5 presents differences in novice teachers' homophily measures relative to teachers in their sixth or later year of teaching. The models with teacher demographic controls illustrate homophily gaps by experience net of the fact that teachers of different races or genders will, by construction, enjoy different degrees of homophily given the composition of the students and other staff in the district.

[5] Controlling directly for their students' prior achievement largely explains lower variation in student prior achievement for elementary and secondary teachers, though not for special education teachers. This may indicate that this lower variation is in part an artifact of floor effects on the tests used here.

As was the case with instructional load, we observe overall homophily gaps favoring veteran teachers for all three types of teacher (column 1), though those gaps are somewhat larger among special education teachers than among elementary teachers, and substantially larger among secondary teachers, amounting to as much as 19 percent of a standard deviation for the least experienced secondary teachers. Also as was the case with instructional load, these gaps are generally smaller among more-experienced novices, and in fact experience gaps in overall homophily scores are only apparent for elementary teachers between veterans and teachers in their first year. This is consistent with both the theory and prior work discussed above suggesting that teachers will be more likely to seek out – or less likely to leave – placements where they enjoy greater congruence with students or other staff. Controlling directly for teacher gender and race (column 2) increases some gaps while shrinking and reversing others. Novice teachers thus appear to enjoy considerably less homophily on the job even net of their own race and gender, particularly among secondary and special education teachers.

Differences in overall homophily scores are driven to a large extent by our racial homophily measures, and novice elementary and secondary teachers in particular experience relatively low rates of racial homophily compared to veterans. For example, first year elementary and secondary teachers have roughly 30 percent lower odds than veterans of having a same-race administrator (column 3) and have approximately six percentage points fewer same-race teachers (column 5) at their school sites. First year elementary teachers also have nearly nine percentage points fewer same-race students (column 7), though this gap is explicable to a large extent by the schools in which teachers work (column 8) or teachers' race (column 9). When novice elementary teachers are compared to teachers in the same school in the same year and controlling for teacher race (column 10), they have only modestly – generally less than two

percentage points – lower shares of same-race students than veterans, suggestive of only minor differences in within-school sorting to same-race students between more and less experienced teachers.

Racial homophily gaps for special education teachers are more muted, and in fact the largest racial homophily gaps favor novice special education teachers, who have 11 to 13 percentage points more same-race students than their veteran counterparts. This reflects the fact that during this period in LAUSD, where the large majority of students are Hispanic, special education teachers in their first five years are 18 percentage points more likely to be Hispanic than veterans. While controlling for teachers' own race does not in general have a large effect on observed racial homophily gaps, it makes those gaps less favorable toward novices among special education teachers, and more like those of elementary and secondary teachers.

Experience gaps in gender homophily are generally smaller in magnitude, reflecting more even distributions of students and staff by gender, and novice elementary teachers enjoy slightly more gender homophily with their coworkers than do veterans. Novice secondary teachers have roughly 30 percent lower odds than veterans of having a same-gender administrator (columns 11 and 12) and new special education teachers have smaller shares of same-gender students (column 15), though this is in part due to the schools in which they teach and gender differences between novice and veteran special education teachers (columns 16 to 18).

Colleague qualifications. As shown in Table 6, the colleague qualification measures considered here point to weaker professional communities for newer teachers. Compared to veteran teachers, teachers in their first year have composite colleague qualification scores (column 1) that are at least 21 percent of a standard deviation lower than those of veterans, and

even teachers in their third through fifth years have scores that are as much as 27 percent of a standard deviation lower at the secondary level.

These overall gaps reflect experience gaps across many individual colleague quality measures. Compared to teachers with six or more years of experience, newer teachers – whether elementary, secondary, or special education – work with less experienced administrators (column 8) and teachers who are less experienced (column 2), and novice secondary and special education teachers' colleagues are less effective as measured by VAM (column 7). Even other measures, where differences are relatively small, suggest weaker colleague qualifications for novice teachers who, relative to veterans, work with slightly smaller shares of teachers who are National Board certified (column 4) and, at the elementary level, who received satisfactory evaluation ratings in the previous year (column 5). In a few cases gaps favor novice teachers; relative to veterans, early career elementary teachers have slightly larger shares of same-site teachers with a graduate degree (column 3), and all three types of novice teacher appear to work with colleagues with slightly higher attendance rates (column 6). Nevertheless, these differences are small and are exceptions to the general pattern of colleague qualification gaps favoring veterans.

Summary. In many respects these results are consistent with prior work finding that novice teachers are more likely than veterans to work with disadvantaged and lower-achieving students, aspects of what we consider instructional load. We also find similar gaps along the other two dimensions of teachers' contexts – homophily and colleague qualifications – that we define using administrative (rather than survey) data. Across a range of measures, then, early career teachers – whether in elementary, secondary, or special education roles – appear to face more difficult or less attractive working conditions than their more-experienced colleagues.

These gaps are concerning but, perhaps just as importantly, we also find evidence suggesting that novice teachers do not necessarily experience similar gaps along every measure of every dimension. One possible explanation is that collective bargaining agreements governing teachers' assignments can exacerbate some gaps (e.g., by allowing more veteran teachers to more easily transfer into placements with fewer low-income students or students of color; Anzia & Moe, 2014) while mitigating others (e.g., by imposing limits on class sizes or assigned course preparations). For example, in addition to setting class size limits, the contract in effect in LAUSD during this time requires that students must be distributed across classes in an "equitable and educationally sound" manner. However, our data do not allow us to determine the mechanisms giving rise to these gaps; understanding why novice teachers experience gaps along some dimensions but not others remains an important area for future work.

RQ2: Are the contexts into which new teachers are placed associated with their effectiveness, development, and retention?

Our final three tables present results of regressions predicting value-added (Table 7), evaluation, attendance (Table 8) and mobility (Table 9) outcomes for teachers using teachers' standardized composite context scores. Because the survey data we use to construct our professional culture measure were typically only released for schools with more than 11 responses and are available in only some years of our panel, including those measures in the model substantially shrinks our sample. We therefore present results for our other three composite measures with and without that sample restriction, though the former are necessarily less precise.

Value-added. As shown in Table 7, teacher VAMs appear most closely related to the composite measure of instructional load; a standard deviation increase in instructional load is

associated in both subjects with VAMs that are at least 26 (18) percent of a teacher-level standard deviation lower in ELA (math) for novices, even in the presence of a teacher fixed effect (columns 2, 6, 8, and 12). The fact that these coefficients tend to shrink somewhat in the presence of a teacher fixed effect (e.g., comparing columns 1 and 2 or columns 7 and 8) suggests that, as observed in prior work (e.g., Goldhaber et al., 2015), there is some sorting of less effective teachers to students with greater need, itself a potential cause for concern. However, even these smaller coefficients are of practical significance, amounting to lower average student achievement of five percent of a standard deviation in both subjects, and are similar in magnitude to the within-teacher effectiveness gains we observe for new ELA teachers between their first and fourth years, or for math teachers between their first and third years. Coefficients on the veteran interaction term in the teacher fixed effect models indicate that these relationships may be smaller in magnitude for veterans, but this varies across subjects and samples. Regardless, it appears that novice teachers' relatively high instructional loads may have genuine implications for their effectiveness.

Our other composite context measures are not consistently associated with teacher VAMs, and relationships often shrink in magnitude and significance (and sometimes switch signs) when controlling for teacher fixed effects or estimating off of only the survey sample. This suggests that the measures that we include in these composites may not be important to teachers' instructional effectiveness, or that their importance varies across the different statewide standardized testing regimes in place before and during the years in which the survey was administered (see Online Appendix B).

Evaluation ratings. Table 8 shows that instructional load and professional culture are the most consistent predictors of teachers' evaluation outcomes. Even after controlling for teacher

fixed effects, a standard deviation improvement in either measure is associated with average EDST scores that are higher by at least five percent of a standard deviation for novices (columns 7 and 11), or roughly 20 percent of the gains in EDST ratings we observe for teachers between their first and second years, and these relationships are not significantly different for veterans. Improvements in instructional load are likewise associated with lower odds of unsatisfactory evaluation ratings, as are improvements in professional culture for veterans (columns 1 and 5) but controlling for teacher fixed effects is more difficult in this case as few teachers receive multiple, and different, ratings. Including a teacher fixed effect for this outcome (column 2) thus reduces our effective sample by more than 95 percent. Other context measures are not consistently related to teachers' evaluation outcomes across samples and specifications, and the fact that coefficients typically shrink in teacher fixed effect specifications suggest again that unobserved differences between teachers can be an important source of bias when relating teachers' working conditions to their outcomes.

Attendance. When predicting teacher attendance, instructional load is again the most consistent predictor of teachers' outcomes. In models with teacher fixed effects (columns 13 and 17) a standard deviation increase in instructional load is associated with novice teachers missing an additional 0.17 to 0.19 percentage points of their scheduled work hours, or roughly one-third of a day in a typical year. The coefficient on colleague qualifications is similar in magnitude in the sample of teachers for whom we have survey data, but only for novices, and the relationship is smaller and insignificant in our full sample of teachers. Other context measures do not significantly predict teacher attendance.

Contrasting models with and without teacher fixed effects again suggests that unobserved differences between teachers are a significant source of bias. In the absence of

teacher fixed effects, for example, homophily and professional culture are significant predictors of teachers' attendance, and coefficients on the interaction terms for instructional load and homophily indicate significantly different relationships for novices and veterans (columns 12, 14, and 16). Including teacher fixed effects causes all of those relationships to lose significance.

Mobility. Though we are unable to control for teacher fixed effects in our multinomial logistic models of teacher mobility, the average marginal probabilities presented in Table 9 are mostly in the expected direction and consistent with prior work. As shown in columns 1 through 4, improvements in instructional load, homophily, colleague qualifications, and professional culture all uniquely predict higher probabilities that a teacher will stay in her school after a given year. Depending on the sample and specification, a standard deviation improvement in each of those measures is associated with an average increase in the probability of a veteran teacher staying in her school of between 0.7 and 1.5 percentage points, or between 0.3 and 2.5 percentage points for a novice, though for novices in a few cases these relationships are not statistically significant. Back-of-the-envelope estimates based on these results suggest that the average context gaps discussed above explain roughly one-tenth of the 10 percentage point within-school retention rate gap between novice and veteran teachers in LAUSD during this time (approximately 80 percent vs. 90 percent).

Estimated increases in retention are due to decreases in both school switching (columns 5 to 8) and district leaving (columns 9 to 12), but the relative magnitude of each varies between novices and veterans. In most cases, higher probabilities of veteran teachers' retention in their schools are due primarily to reductions in school switching, while for novices they are due primarily to lower odds of leaving the district. This may reflect the relatively greater ability of veterans to secure positions in other district schools when they are dissatisfied with their

placements (e.g., due to accumulated seniority rights); novices may have little recourse other than to leave the district altogether.

Summary. Collectively, these results provide further evidence that the features of teachers' placements are associated with their outcomes. Even net of teacher demographics and experience levels, and making within-year comparisons of teachers, composite measures of teachers' contexts are meaningfully predictive of their attendance, VAM, evaluation, and mobility outcomes. In some cases, estimates are demonstrably sensitive to controlling for unobserved heterogeneity between teachers, highlighting difficulties in interpreting correlational relationships between teachers' working conditions and their outcomes. However, in other cases estimates persist even net of teacher fixed effects, particularly in the case of instructional load. These results provide further evidence that teachers' contexts can be characterized more finely than has often been done in the past – e.g., as "challenging," "hard to staff," or "facing difficult working conditions." And finer-grained characterizations may also have utility; the four dimensions defined here are not only largely orthogonal to each other, but are also differentially predictive of teacher outcomes, suggesting that they may each have distinct explanatory power for patterns of new teacher development and retention.

Discussion

Given the central importance of teachers in the effectiveness of schools, the way novice teachers sort into schools and classrooms is likely to have important implications not only for the quality of teacher to which different kinds of student are exposed, but also for the professional trajectories of beginning teachers. Accordingly, a growing body of literature documents these patterns of new teacher sorting and links characteristics of novice teachers' contexts to their outcomes. Nevertheless, data limitations often hamper researchers' ability to characterize these

contexts comprehensively, and perhaps as a result these characterizations have often been unidimensional and poorly defined (e.g., “challenging”), or defined in ways that are difficult to operationalize in practice (e.g., because definitions emerge from survey data).

We contribute to this literature in several ways. First, the quantity and detail of these data allow us to characterize novice teachers’ contexts across more dimensions than in much of the earlier literature. Consistent with prior research, we find that novice teachers are indeed more likely than their more veteran counterparts to sort into schools, and then again into classrooms, with lower levels of student achievement, greater disciplinary issues, and larger shares of English language learners; that is, into placements that are sometimes characterized as “challenging.” However, we also extend prior work by showing that relatively novice teachers are likely to sort into schools where they will enjoy less racial and gender homophily with the students and other staff, and into schools with less experienced colleagues who hold weaker credentials and make smaller contributions to student achievement.

Second, we define four dimensions of teacher context: namely, as having greater or lesser instructional load, homophily, and colleague qualifications, or stronger or weaker professional culture. When measured on a composite basis, these dimensions are largely orthogonal to one another, suggesting that characterizing novice teachers’ contexts and working conditions may be more complex than has been previously assumed. Third, these composite measures of contexts are often significantly, uniquely, and differentially predictive of beginning teachers’ outcomes. Thus, we provide evidence that these dimensions not only each capture distinct information, but information that is important for the support of the new teacher workforce.

Because, like most similar literature, we do not observe plausibly exogenous variation in these contextual features, these results do not have a clear causal interpretation. In particular, both our individual measures of teachers' contexts and our composite measures may simply serve as proxies for other, unobserved teacher, school, and student attributes that are correlated with both teacher experience and teachers' outcomes. Though our results are often robust to the inclusion of teacher fixed effects, particularly in the case of instructional load, other estimates are not, highlighting the need to interpret correlational relationships between teachers' working conditions and outcomes with caution. We are also far from knowing the most useful, accurate, and comprehensive ways in which to characterize teachers' professional contexts. For example, it is neither obvious that "colleague qualifications" is an optimal dimension by which to characterize teachers' placements, nor clear that the measures that dimension comprises are correctly chosen.

Nonetheless, these findings will be of interest to district and state education leaders concerned with the retention and development of their novice teachers. District administrators hoping to implement staffing policies that support their novice teachers may need to attend to a wider range of contextual features than has typically been considered in the past, such as the quality and demographic composition of a teachers' coworkers, in addition to ensuring that novice teachers are not assigned disproportionately to the highest-need students. It is also of practical significance that our composite measures of teachers' contexts are only weakly correlated with one another. This suggests that policymakers cannot assume that a teaching placement that is desirable in one respect will be desirable in general; a placement that imposes low levels of instructional load on a teacher may at the same time offer relatively little homophily or colleagues with relatively weak qualifications. Administrators may therefore need

to balance competing considerations in their novice teacher placements and may need to think more carefully about the trade-offs entailed by their staffing practices. Given the complexity of teaching itself, it should perhaps not surprise us to learn that placing novice teachers into suitable classrooms is subtle and complex as well.

References

Allensworth, E., & Easton, J. Q. (2007). *What Matters for Staying On-Track and Graduating in Chicago Public High Schools* (Research Report). Consortium on Chicago School Research: University of Chicago.

Anzia, S. F., & Moe, T. M. (2014). Collective Bargaining, Transfer Rights, and Disadvantaged Schools. *Educational Evaluation and Policy Analysis*, *36*(1), 83–111.

Boyd, D., Grossman, P., Ing, M., Lankford, H., Loeb, S., & Wyckoff, J. (2011). The Influence of School Administrators on Teacher Retention Decisions. *American Educational Research Journal*, *48*(2), 303–333.

Boyd, D., Lankford, H., Loeb, S., Ronfeldt, M., & Wyckoff, J. (2011). The role of teacher quality in retention and hiring: Using applications to transfer to uncover preferences of teachers and schools. *Journal of Policy Analysis and Management*, *30*(1), 88–110.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). Explaining the Short Careers of High-Achieving Teachers in Schools with Low-Performing Students. *American Economic Review*, *95*(2), 166–171.

Burnette, D. (2017, May 30). States Struggle to Define "Ineffective Teachers" Under ESSA. *Education Week*, pp. 1–17.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review*, *104*(9), 2633–2679.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. (2005). Who teaches whom? Race and the distribution of novice teachers. *Economics of Education Review*, *24*(4), 377–392.

Dee, T. S. (2004). Teachers, Race, and Student Achievement in a Randomized Experiment. *The Review of Economics and Statistics*, *86*(1), 195–210.

Dee, T. S., & Goldhaber, D. (2017). *Understanding and Addressing Teacher Shortages in the United States* (The Hamilton Project). Washington, D.C.: The Brookings Institution.

Dee, T. S., & Wyckoff, J. (2015). Incentives, Selection, and Teacher Performance: Evidence from IMPACT. *Journal of Policy Analysis and Management*, *34*(2), 267–297.

Donaldson, M. L., & Johnson, S. M. (2010). The Price of Misassignment: The Role of Teaching Assignments in Teach For America Teachers' Exit From Low-Income Schools and the Teaching Profession. *Educational Evaluation and Policy Analysis*, *32*(2), 299–323.

Downey, D. B., & Pribesh, S. (2004). When Race Matters: Teachers' Evaluations of Students' Classroom Behavior. *Sociology of Education*, *77*(4), 267–282.

Egalite, A. J., Kisida, B., & Winters, M. A. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review*, *45*, 44–52.

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

Feng, L., & Sass, T. R. (2017). Teacher Quality and Teacher Mobility. *Education Finance and Policy*, *12*(3), 396–418.

Figlio, D. N. (2007). Boys Named Sue: Disruptive Children and Their Peers. *Education Finance and Policy*, *2*(4), 376–394.

Gershenson, S., Holt, S. B., & Papageorge, N. W. (2016). Who Believes in Me? The Effect of Student-Teacher Demographic Match on Teacher Expectations. *Economics of Education Review*, *52*, 209–224.

Goldhaber, D., Lavery, L., & Theobald, R. (2015). Uneven Playing Field? Assessing the Teacher Quality Gap Between Advantaged and Disadvantaged Students. *Educational Researcher*, *44*(5), 293–307.

Goldhaber, D., Lavery, L., & Theobald, R. (2016). Inconvenient Truth? Do Collective Bargaining Agreements Help Explain the Mobility of Teachers within School Districts? *Journal of Policy Analysis and Management*, *35*(4), 848–880.

Gottfried, M. A. (2014). Chronic Absenteeism and Its Effects on Students' Academic and Socioemotional Outcomes. *Journal of Education for Students Placed at Risk*, *19*(2), 53–75.

Grissom, J. A. (2011). Can Good Principals Keep Teachers in Disadvantaged Schools? Linking Principal Effectiveness to Teacher Satisfaction and Turnover in Hard-to-Staff Environments. *Teachers College Record*, *113*(11), 2552–2585.

Grissom, J. A., & Keiser, L. R. (2011). A supervisor like me: Race, representation, and the satisfaction and turnover decisions of public sector employees. *Journal of Policy Analysis and Management*, *30*(3), 557–580.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Why Public Schools Lose Teachers. *Journal of Human Resources*, *XXXIX*(2), 326–354.

Hanushek, E. A., & Rivkin, S. G. (2007). Pay, working conditions, and teacher quality. *The Future of Children*, *17*(1), 69–86.

Hanushek, E. A., & Rivkin, S. G. (2012). The Distribution of Teacher Quality and Implications for Policy. *Annual Review of Economics*, *4*(1), 131–157.

Jackson, C. K., & Bruegmann, E. (2009). Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers. *American Economic Journal: Applied Economics*, *1*(4), 85–108.

Johnson, S. M., Kraft, M. A., & Papay, J. P. (2012). How Context Matters in High-Need Schools: The Effects of Teachers' Working Conditions on Their Professional Satisfaction and Their Students' Achievement. *Teachers College Record*, *114*(10).

Kalogrides, D., Loeb, S., & Béteille, T. (2013). Systematic Sorting Teacher Characteristics and Class Assignments. *Sociology of Education*, *86*(2), 103–123.

Keigher, A. (2010). *Teacher Attrition and Mobility: Results from the 2008-09 Teacher Follow-up Survey* (No. NCES 2010-353). Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

Kelly, S. (2004). Are Teachers Tracked? On what Basis and with what Consequences. *Social Psychology of Education*, *7*(1), 55–72.

Koedel, C., Mihaly, K., & Rockoff, J. E. (2015). Value-added modeling: A review. *Economics of Education Review*, *47*, 180–195.

Kraft, M. A. (2017). Teacher Effects on Complex Cognitive Skills and Social-Emotional Competencies. *Journal of Human Resources*, 0916-8265R3.

Kraft, M. A., Marinell, W. H., & Yee, D. S.-W. (2016). School Organizational Contexts, Teacher Turnover, and Student Achievement Evidence From Panel Data. *American Educational Research Journal*, *53*(5), 1411–1449.

Kraft, M. A., & Papay, J. P. (2014). Can Professional Environments in Schools Promote Teacher Development? Explaining Heterogeneity in Returns to Teaching Experience. *Educational Evaluation and Policy Analysis*, *36*(4), 476–500.

Kukla-Acevedo, S. (2009). Leavers, Movers, and Stayers: The Role of Workplace Conditions in Teacher Mobility Decisions. *The Journal of Educational Research*, *102*(6), 443–452.

Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis. *Educational Evaluation and Policy Analysis*, *24*(1), 37–62.

Maranto, R., & Shuls, J. V. (2012). How Do We Get Them on the Farm? Efforts to Improve Rural Teacher Recruitment and Retention in Arkansas. *Rural Educator*, *34*(1).

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, *27*(1), 415–444.

MetLife Foundation. (2009). *The MetLife survey of the American teacher: Collaborating for student success*.

Milanowski, A. T., & Odden, A. R. (2007). *A New Approach to the Cost of Teacher Turnover*. School Finance Redesign Project Working Paper 13.

Neild, R. C., & Farley-Ripple, E. (2008). Within-School Variation in Teacher Quality: The Case of Ninth Grade. *American Journal of Education*, *114*(3), 271–305.

Ost, B. (2014). How Do Teachers Improve? The Relative Importance of Specific and General Human Capital. *American Economic Journal: Applied Economics*, *6*(2), 127–151.

Ost, B., & Schiman, J. C. (2017). Workload and Teacher Absence. *Economics of Education Review*, *57*, 20–30.

Podolsky, A., Kini, T., Bishop, J., & Darling-Hammond, L. (2016). *Solving the Teacher Shortage: How to Attract and Retain Excellent Educators*. Palo Alto, CA: Learning Policy Institute.

Ronfeldt, M., Farmer, S. O., McQueen, K., & Grissom, J. A. (2015). Teacher Collaboration in Instructional Teams and Student Achievement. *American Educational Research Journal*, *52*(3), 475–514.

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How Teacher Turnover Harms Student Achievement. *American Educational Research Journal*, *50*(1), 4–36.

Rothstein, J. (2009). Student Sorting and Bias in Value-Added Estimation: Selection on Observables and Unobservables. *Education Finance and Policy*, *4*(4), 537–571.

Scafidi, B., Sjoquist, D. L., & Stinebrickner, T. R. (2007). Race, poverty, and teacher mobility. *Economics of Education Review*, *26*(2), 145–159.

Strunk, K. O., & Robinson, J. P. (2006). Oh, Won't You Stay: A Multilevel Analysis of the Difficulties in Retaining Qualified Teachers. *Peabody Journal of Education*, *81*(4), 65–94.

Will, M. (2017, November 21). Here Are Six Strategies for States to Build Stronger Teacher Pipelines. *Education Week*.

Wright, A., Gottfried, M. A., & Le, V.-N. (2017). A Kindergarten Teacher Like Me: The Role of Student-Teacher Race in Social-Emotional Development. *American Educational Research Journal*, *54*(1_suppl), 78S-101S.

Tables

Table 1 – Summary Statistics: School Characteristics

	N	Mean	SD	Min	Max
Teachers	7371	35.82	26.76	2	238
Administrators					
Principals	7371	1.00	0.09	0	3
Assistant Principals	7371	0.77	1.19	0	6
% Female	7371	64.40	41.26	0.0	100
% Native American	7371	0.21	4.05	0.0	100
% Asian	7371	7.05	22.30	0.0	100
% Black	7371	20.36	35.80	0.0	100
% Filipino	7371	0.91	8.02	0.0	100
% Hispanic	7371	36.19	42.61	0.0	100
% Pacific Islander	7371	0.15	3.53	0.0	100
% White	7371	34.42	42.17	0.0	100
% with Doctorate	7371	6.19	20.86	0.0	100
Avg. Years Admin. Experience	7334	9.04	4.90	0.0	33.6
Student Characteristics					
Students	7371	755.38	619.69	29	4478
% Female	7371	48.47	5.30	17.0	100
% Native American	7371	0.33	0.44	0.0	9.6
% Asian	7371	4.46	7.58	0.0	73.7
% Black	7371	10.24	15.54	0.0	95.3
% Filipino	7371	1.59	3.42	0.0	46.3
% Hispanic	7371	72.98	25.63	1.1	100
% Pacific Islander	7371	0.38	0.98	0.0	13.6
% White	7371	10.02	17.20	0.0	89.0
% FRL	7371	76.38	23.49	0.5	100
% EL	7371	29.20	17.75	0.0	92.0
% SPED	7371	12.22	11.73	0.0	100
Prior Math Achievement	7097	0.14	0.56	-2.3	2.0
Prior ELA Achievement	7096	0.02	0.46	-3.2	1.7
Prior Social Studies Achievement	1258	-0.27	0.48	-2.1	1.4
Prior Science Achievement	2708	-0.08	0.51	-2.8	1.4
Prior Overall GPA	7326	-0.02	0.42	-4.4	1.6
Prior Math GPA	7302	-0.02	0.38	-3.2	1.9
Prior ELA GPA	7301	-0.03	0.40	-3.5	1.7
Prior Social Studies GPA	7294	-0.03	0.39	-4.5	1.3
Prior Science GPA	7300	-0.03	0.39	-4.5	1.4
Prior Physical Education GPA	7326	-0.04	0.40	-4.8	2.2
Prior Art GPA	7302	-0.03	0.38	-5.5	1.5
% Days Absent Last Year	7352	5.08	4.10	0.8	72.8
% Suspended Last Year	7355	2.12	4.12	0.0	53.5
Survey Measures (Percentage of Staff Responding Agree/Strongly Agree or Always/Often)					
PD Meets Students' Needs	2619	46.11	25.00	0.0	100
PD Appropriate for My Experience	2621	43.15	22.75	0.0	100
Performance Review Helps Improvement	2617	39.89	21.31	0.0	100
Teachers have Sufficient Autonomy	2620	64.76	23.20	0.0	100
Parents are Partners	2641	49.27	25.73	0.0	100
Trust Amongst Staff	2552	56.77	23.42	0.0	100
Comfortable Talking with Leaders	2524	59.42	21.36	3.0	100
There is Team Alignment	2613	66.73	17.82	0.0	100
Leaders Visit to Observe	2612	52.94	21.55	0.0	100
Leaders Provide Useful Feedback	2610	50.57	19.04	0.0	100

Note. Data in this table combines annual observations of 808 unique schools from 2007-8 through 2016-17.

Table 2 – Summary Statistics: Teacher Characteristics

	Teachers in Years 1-5 (Novice)					Teachers in Year 6+ (Veteran)				
	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max
Elementary	31549	0.31	0.46	0	1	228702	0.55	0.50	0	1
Secondary	31549	0.42	0.49	0	1	228702	0.32	0.46	0	1
Special Education	31549	0.27	0.44	0	1	228702	0.14	0.34	0	1
Female	31547	0.72	0.45	0	1	228702	0.69	0.46	0	1
Native American	29944	0.01	0.09	0	1	228163	0.01	0.08	0	1
Asian	29944	0.13	0.33	0	1	228163	0.09	0.28	0	1
Black	29944	0.09	0.28	0	1	228163	0.11	0.32	0	1
Filipino	29944	0.03	0.16	0	1	228163	0.03	0.17	0	1
Hispanic	29944	0.38	0.49	0	1	228163	0.35	0.48	0	1
Pacific Islander	29944	0.00	0.05	0	1	228163	0.00	0.05	0	1
White	29944	0.37	0.48	0	1	228163	0.41	0.49	0	1
Has MA or Doctorate	31330	0.29	0.45	0	1	228123	0.40	0.49	0	1
Experience										
First Year	31549	0.18	0.38	0	1	228702	0	0	0	0
Second Year	31549	0.19	0.39	0	1	228702	0	0	0	0
Years 3-5	31549	0.63	0.48	0	1	228702	0	0	0	0
Year 6 or More	31549	0	0	0	0	228702	1	0	1	1
Outcomes										
Attendance Rate	11030	96.67	4.86	0	100	132299	95.93	6.92	0	100
Math VAM	6590	0.01	0.95	-3.39	5.45	50832	0.01	1.01	-3.51	5.80
ELA VAM	6667	-0.07	0.94	-3.12	4.65	53337	0.02	1.01	-4.11	6.98
Below Standard Evaluation	19062	0.02	0.12	0	1	74771	0.03	0.17	0	1
Average EDST Rating	5816	-0.01	0.85	-4.57	2.23	18971	0.00	1.04	-4.97	2.23
Switch School	30868	0.12	0.33	0	1	228425	0.07	0.26	0	1
Leave District	30868	0.08	0.27	0	1	228425	0.04	0.21	0	1

Note. Combines annual observations of 40,879 unique teachers from 2007-8 through 2016-17.

Table 3 – Summary Statistics: Teacher Context Measures

	Teachers in Years 1-5					Teachers in Years 6+				
	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max
Instructional Load										
Prior Achievement (Average of Tests)	23393	-0.31	0.60	-2.96	2.38	144039	-0.05	0.65	-3.03	2.65
Prior Achievement (Teacher's Subject)	7449	-0.24	0.45	-2.58	2.32	33587	-0.07	0.54	-2.60	3.04
SD of Achieve. (math and ELA)	23008	0.69	0.18	0	2.34	140545	0.74	0.17	0	2.35
SD of Achieve. (Teachers' Subject)	7417	0.72	0.18	0	2.14	33483	0.76	0.18	0	2.21
Prior GPA (Overall)	27260	-0.21	0.69	-4.62	2.71	191835	-0.07	0.67	-4.64	2.71
Prior GPA (Teacher's Subject)	10465	-0.03	0.35	-1.84	1.40	52130	0.03	0.38	-1.78	1.77
% SPED-eligible	22285	7.57	6.38	0	100	185787	7.33	6.87	0	100
% EL	28921	34.46	27.00	0	100	207474	29.69	29.33	0	100
% FRL-eligible	28921	76.55	21.69	0	100	207471	77.18	23.69	0	100
% Suspended Last Year	28353	4.97	7.20	0	100	200315	2.49	5.14	0	100
Avg. Absence Rate Last Year	28335	5.39	3.27	0	84.17	200136	4.72	3.61	0	100
Class Size	28934	23.00	9.29	3.08	77.25	207022	24.38	8.33	3.10	78.50
Preparations (Secondary)	12262	5.30	1.95	1	14	63500	5.49	2.13	1	15
Instructional Load measures observed	31549	8.99	3.47	0	13	228702	8.26	3.39	0	13
Instructional Load Score	29092	0.28	1	-3.55	14.96	208863	-0.04	0.99	-5.38	15.36
Homophily										
Any Same-Race Admin.	29944	0.61	0.49	0	1	228163	0.58	0.49	0	1
% Same-Race Teachers	29944	36.09	21.50	0	100	228163	40.20	23.03	0	100
% Same-Race Students	27484	38.94	40.40	0	100	207024	40.14	40.15	0	100
Any Same-Gender Admin.	31547	0.86	0.35	0	1	228702	0.79	0.41	0	1
% of Same-Gender Teachers	31547	60.81	19.98	0	100	228702	62.40	22.99	0	100
% Same-Gender Students	28919	47.87	14.02	0	100	207474	49.74	11.82	0	100
Homophily measures observed	31549	5.69	0.84	2	6	228702	5.81	0.60	2	6
Homophily Score	31549	-0.11	0.97	-4.51	3.83	228702	0.01	1	-4.53	3.76
Colleague Qualifications										
Other Teachers:										
% with 10+ Years Experience	31549	59.68	18.62	0	100	228702	74.12	16.30	0	100
% with Graduate Degree	31549	37.85	10.63	0	100	228702	38.44	11.17	0	100
% NBC	31549	2.94	3.87	0	50	228702	4.14	4.77	0	100
% Unsatisfactory Last Year	30871	2.16	6.53	0	100	222106	2.33	7.58	0	100
Avg. Prior Attendance Rate	9485	96.45	1.29	60.38	100	108337	96.37	1.25	47.60	100
Avg. Prior VAM	28278	-0.05	0.46	-2.59	2.42	198196	0	0.46	-3.38	2.42
Avg. Years Admin. Experience	31500	7.73	3.86	0.01	31.69	228050	9.07	4.46	0.01	33.63
% Admin. with Doctorate	31549	0.07	0.17	0	1	228702	0.07	0.20	0	1
Colleague Qual. measures observed	31549	7.17	0.47	5	8	228702	7.31	0.54	5	8
Colleague Qualifications Score	31549	-0.28	1.04	-12.32	9.41	228702	0.04	0.99	-12.83	9.66
Professional Culture (Percentage of Staff Responding Agree/Strongly Agree or Always/Often)										
PD Meets Students' Needs	7568	39.08	22.75	0	100	77963	43.91	24.03	0	100
PD Appropriate for My Experience	7562	37.16	20.66	0	100	77990	41	21.73	0	100
Perform. Review Helps Improvement	7564	34.74	19.67	0	96.00	77897	38.22	20.53	0	100
Teachers have Sufficient Autonomy	7559	59.04	23.04	3.00	100	77973	63.53	22.53	0	100
Parents are Partners	7596	39.74	23.14	0	100	78491	46.62	24.60	0	100
Trust Amongst Staff	7131	49.51	22.04	0	100	73604	54.25	22.84	0	100
Comfortable Talking with Leaders	7064	52.68	20.52	3.00	100	72514	56.86	20.93	3.00	100
There is Team Alignment	7566	63.45	17.00	5.00	100	77860	65.09	17.12	0	100
Leaders Visit to Observe	7534	47.97	21.26	0	100	77748	48.77	21.27	0	100
Leaders Provide Useful Feedback	7537	46.11	18.21	0	100	77645	47.23	18.51	0	100
Prof. Culture measures observed	9485	7.87	3.91	0	10	108338	7.10	4.39	0	10
Professional Culture Score	7728	-0.27	0.96	-3.03	3.09	79723	-0.20	0.96	-3.03	3.38

Note. Combines annual observations of 40,879 unique teachers from 2007-8 through 2016-17. See Appendix for professional culture survey questions.

Table 4 – New Teachers' Instructional Loads Compared to Teachers in Year 6+

		Prior Achievement		SD of Prior Achievement		Prior GPA		Teacher's Students						
	IL Score	Average of Tests	Teacher's Subject	Math/ ELA	Teacher's Subject	Overall	Teacher's Subject	% SPED	% EL	% FRL	% Suspended	Absence Rate	Class Size	Preps
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Experience Level (Reference Group = Teachers with 6 or more years of experience)														
First	0.13***	-0.19***		-0.03***		-0.08***		-0.02	0.43	1.37	0.79***	0.15*	0.52**	
Year	(0.03)	(0.02)		(0.01)		(0.02)		(0.23)	(0.89)	(0.87)	(0.15)	(0.07)	(0.18)	
Second	0.13***	-0.19***		-0.03***		-0.05**		0.26	1.26	1.21	0.62***	0.22*	0.86***	
Year	(0.03)	(0.02)		(0.01)		(0.02)		(0.18)	(0.84)	(0.84)	(0.15)	(0.09)	(0.18)	
Years	0.06***	-0.10***		-0.02***		-0.04***		0.11	0.43	-0.59	0.37***	0.06	0.85***	
3-5	(0.02)	(0.01)		(0.00)		(0.01)		(0.10)	(0.50)	(0.49)	(0.07)	(0.05)	(0.08)	
School-by-Year Fixed Effects														
First	0.09***	-0.12***		-0.02***		-0.08***		0.15	0.52	0.80**	0.42**	0.07	0.30*	
Year	(0.02)	(0.02)		(0.01)		(0.02)		(0.24)	(0.74)	(0.29)	(0.14)	(0.07)	(0.12)	
Second	0.08***	-0.13***		-0.02**		-0.06***		0.43*	1.13	0.39	0.11	0.16*	0.50***	
Year	(0.02)	(0.02)		(0.01)		(0.02)		(0.19)	(0.73)	(0.27)	(0.10)	(0.07)	(0.11)	
Years	0.06***	-0.07***		-0.01***		-0.05***		0.35***	0.49	0.12	0.06	0.04	0.59***	
3-5	(0.01)	(0.01)		(0.00)		(0.01)		(0.11)	(0.38)	(0.15)	(0.06)	(0.04)	(0.06)	
Teachers	20000	14617		13869		19340		19979	19979	19979	19660	19649	19952	
N	126283	64534		61535		111304		125522	125522	125521	117172	117044	126048	
Secondary Teachers														
First	0.26***	-0.22***	-0.27***	-0.04***	-0.07***	-0.19***	-0.16***	1.04***	7.13***	5.65***	1.90***	0.22*	-1.66***	-0.49***
Year	(0.02)	(0.01)	(0.02)	(0.00)	(0.01)	(0.01)	(0.01)	(0.13)	(0.48)	(0.58)	(0.22)	(0.09)	(0.23)	(0.06)
Second	0.19***	-0.17***	-0.21***	-0.03***	-0.04***	-0.10***	-0.07***	0.91***	4.89***	5.32***	1.20***	0.05	-0.95***	-0.17***
Year	(0.02)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.12)	(0.44)	(0.53)	(0.17)	(0.08)	(0.20)	(0.05)
Years	0.12***	-0.10***	-0.14***	-0.02***	-0.02***	-0.06***	-0.04***	0.51***	3.18***	3.95***	0.56***	-0.07	-0.82***	0.02
3-5	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.08)	(0.27)	(0.33)	(0.09)	(0.05)	(0.12)	(0.03)
School-by-Year Fixed Effects														
First	0.11***	-0.09***	-0.13***	-0.02***	-0.04***	-0.13***	-0.11***	0.64***	2.62***	0.37**	1.26***	0.27***	-1.01***	-0.44***
Year	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.11)	(0.35)	(0.14)	(0.11)	(0.03)	(0.15)	(0.04)
Second	0.07***	-0.05***	-0.09***	-0.01***	-0.02***	-0.06***	-0.04***	0.55***	1.16**	0.32*	0.70***	0.11***	-0.37**	-0.10*
Year	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.12)	(0.36)	(0.14)	(0.09)	(0.03)	(0.14)	(0.04)
Years	0.05***	-0.03***	-0.06***	-0.01***	-0.01***	-0.03***	-0.02***	0.30***	0.68***	0.30**	0.43***	0.07***	-0.47***	0.08***
3-5	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.07)	(0.21)	(0.09)	(0.05)	(0.02)	(0.09)	(0.02)
Teachers	14769	14748	8745	14744	8713	14764	11882	14769	14769	14769	14767	14767	14639	13824
N	82560	82433	41036	82338	40900	82538	62595	82550	82550	82550	82552	82552	81475	75762
Special Education Teachers														
First	0.21***	-0.11***		-0.02*		-0.21***			3.39***	2.56***	0.78**	0.04	-0.20	
Year	(0.03)	(0.02)		(0.01)		(0.03)			(0.74)	(0.58)	(0.27)	(0.17)	(0.13)	
Second	0.16***	-0.13***		-0.03**		-0.18***			3.77***	3.40***	0.52+	-0.22+	-0.01	
Year	(0.03)	(0.01)		(0.01)		(0.03)			(0.70)	(0.57)	(0.27)	(0.13)	(0.12)	
Years	0.09***	-0.08***		-0.01**		-0.14***			2.30***	2.07***	0.43*	-0.35***	0.26**	
3-5	(0.02)	(0.01)		(0.00)		(0.02)			(0.45)	(0.39)	(0.19)	(0.08)	(0.08)	
School-by-Year Fixed Effects														
First	0.12***	-0.06**		-0.01		-0.05*			1.07	-0.05	0.49+	0.49***	-0.45**	
Year	(0.03)	(0.02)		(0.01)		(0.02)			(0.67)	(0.43)	(0.29)	(0.14)	(0.15)	
Second	0.10***	-0.09***		-0.01		-0.05*			2.17***	0.67	0.43	0.18	-0.26+	
Year	(0.03)	(0.02)		(0.01)		(0.02)			(0.62)	(0.41)	(0.27)	(0.12)	(0.13)	
Years	0.02	-0.04***		-0.01		-0.01			0.25	-0.45	0.46*	0.01	0.07	
3-5	(0.02)	(0.01)		(0.01)		(0.01)			(0.40)	(0.28)	(0.19)	(0.08)	(0.09)	
Teachers	6086	4868		4713		5813			6069	6068	6065	6058	6027	
N	29112	20465		19680		25253			28323	28321	28944	28875	28433	

Note. Standard errors clustered on school-by-year in parentheses. Includes teachers from 2007-8 through 2016-17. Models without school-by-year fixed effects include year fixed effects.

+ p<.1, * p<.05 ** p<.01 *** p<.001

Table 5 – New Teachers' Homophily Compared to Teachers in Years 6+

			Same-Race								Same-Gender							
	Homophily Score		Any Admin.[a]		% Teachers		% Students				Any Admin.[a]		% Teachers		% Students			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Elementary Teachers																		
First Year	-0.12***	-0.09***	0.70***	0.82**	-6.36***	-3.68***	-8.54***	-3.98**	-2.26**	-1.65*	1.37***	1.23*	3.20***	-0.46	0.02	-0.06	0.11	0.04
	(0.03)	(0.03)	(0.05)	(0.06)	(0.69)	(0.59)	(1.31)	(1.26)	(0.71)	(0.70)	(0.13)	(0.11)	(0.69)	(0.41)	(0.25)	(0.27)	(0.25)	(0.27)
Second Year	-0.02	-0.04+	0.84**	0.95	-3.92***	-2.46***	-1.89+	1.69	-0.86	-0.15	1.30**	1.16+	2.87***	-0.56	-0.07	-0.28	0.02	-0.17
	(0.03)	(0.02)	(0.05)	(0.06)	(0.61)	(0.54)	(1.11)	(1.07)	(0.60)	(0.55)	(0.11)	(0.10)	(0.62)	(0.36)	(0.22)	(0.23)	(0.22)	(0.22)
Years 3-5	0.01	-0.05***	0.84***	0.91**	-3.09***	-2.47***	0.56	2.82***	-1.33***	-0.85**	1.33***	1.19***	3.20***	-0.60**	-0.19	-0.23+	-0.09	-0.13
	(0.02)	(0.01)	(0.03)	(0.03)	(0.35)	(0.31)	(0.64)	(0.60)	(0.36)	(0.33)	(0.07)	(0.06)	(0.35)	(0.21)	(0.12)	(0.13)	(0.12)	(0.13)
Teacher Controls		X		X		X			X	X		X		X			X	X
School-by-Year FE								X		X						X		X
Teachers	20896	20849	20849	20686	20849	20849	19937	19937	19937	19937	20896	20896	20896	20896	19979	19979	19979	19979
Observations	135175	134919	134919	133880	134919	134919	125289	125289	125289	125289	135175	135175	135175	135175	125522	125522	125522	125522
Secondary Teachers																		
First Year	-0.19***	-0.16***	0.67***	0.70***	-5.55***	-3.37***	-1.63	-3.22**	-2.59***	-2.12***	0.66***	0.66***	0.46+	-0.39+	-0.77***	-0.58**	-0.73***	-0.53**
	(0.02)	(0.02)	(0.04)	(0.05)	(0.54)	(0.42)	(1.07)	(1.09)	(0.39)	(0.44)	(0.05)	(0.05)	(0.24)	(0.21)	(0.19)	(0.20)	(0.19)	(0.20)
Second Year	-0.16***	-0.14***	0.70***	0.71***	-4.66***	-2.47***	1.45	-0.37	-2.10***	-1.71***	0.67***	0.67***	0.39+	-0.46*	-0.75***	-0.59**	-0.70***	-0.53**
	(0.02)	(0.02)	(0.04)	(0.04)	(0.49)	(0.38)	(0.94)	(0.94)	(0.36)	(0.38)	(0.06)	(0.06)	(0.21)	(0.19)	(0.17)	(0.18)	(0.17)	(0.18)
Years 3-5	-0.07***	-0.09***	0.82***	0.83***	-3.58***	-1.61***	4.49***	3.00***	-1.14***	-0.74***	0.71***	0.72***	0.49***	-0.23*	-0.29**	-0.19+	-0.25*	-0.15
	(0.01)	(0.01)	(0.03)	(0.03)	(0.29)	(0.21)	(0.54)	(0.55)	(0.20)	(0.20)	(0.04)	(0.04)	(0.13)	(0.11)	(0.10)	(0.11)	(0.10)	(0.11)
Teacher Controls		X		X		X			X	X		X		X			X	X
School-by-Year FE								X		X						X		X
Teachers	15145	14948	14948	14948	14948	14948	14568	14568	14568	14568	15145	15145	15145	15145	14769	14769	14769	14769
Observations	85454	84331	84331	84331	84331	84331	81450	81450	81450	81450	85454	85454	85454	85454	82550	82550	82550	82550
Special Education Teachers																		
First Year	-0.09**	-0.15***	0.97	0.86*	-1.63**	-2.30***	11.42***	10.95***	-1.28*	-0.98	0.93	0.90	1.54**	-0.39	-4.60***	-3.27***	-1.90***	-1.42*
	(0.03)	(0.03)	(0.06)	(0.05)	(0.62)	(0.52)	(1.23)	(1.50)	(0.61)	(0.72)	(0.06)	(0.06)	(0.55)	(0.40)	(0.67)	(0.83)	(0.47)	(0.57)
Second Year	-0.06*	-0.14***	0.97	0.86**	-1.49*	-2.26***	13.38***	12.11***	-0.63	-0.76	1.03	0.99	1.46**	-0.73+	-4.48***	-3.10***	-1.23**	-0.77
	(0.03)	(0.02)	(0.05)	(0.05)	(0.60)	(0.49)	(1.16)	(1.36)	(0.57)	(0.70)	(0.07)	(0.07)	(0.54)	(0.40)	(0.63)	(0.78)	(0.45)	(0.55)
Years 3-5	-0.06***	-0.13***	0.96	0.88***	-1.75***	-2.26***	10.56***	9.97***	-1.44***	-1.00*	1.03	0.99	2.29***	-0.07	-5.22***	-4.02***	-1.82***	-1.56***
	(0.02)	(0.01)	(0.03)	(0.03)	(0.37)	(0.30)	(0.72)	(0.90)	(0.34)	(0.42)	(0.04)	(0.04)	(0.34)	(0.25)	(0.40)	(0.49)	(0.29)	(0.36)
Teacher Controls		X		X		X			X	X		X		X			X	X
School-by-Year FE								X		X						X		X
Teachers	7519	7417	7417	7350	7417	7417	5967	5967	5967	5967	7519	7519	7519	7519	6069	6069	6069	6069
Observations	39622	38855	38857	38492	38857	38857	27769	27769	27769	27769	39620	39620	39620	39620	28321	28321	28321	28321

Note. Standard errors clustered on school-by-year in parentheses. Includes teachers from 2007-8 through 2016-17. Models without school-by-year fixed effects include year fixed effects. Teacher controls include teacher race or gender dummies in race- or gender-homophily models, respectively, or both for overall homophily scores.

[a] Logistic regressions. Coefficients are odds ratios.

+ p<.1, * p<.05 ** p<.01 *** p<.001

Table 6 – New Teachers Colleague Qualifications Compared to Teachers in Years 6+

		Other Teachers						Administrators	
Elementary Teachers	Colleague Qualifications Score (1)	% with 10+ Years Experience (2)	% with Graduate Degree (3)	% NBC (4)	% Below Standard (5)	Avg. Attendance Rate (6)	Avg. VAM (7)	Avg. Years Experience (8)	% with Doctorate (9)
Experience Level (Reference Group = Teachers with 6 or more years of experience)									
First Year	-0.38***	-10.81***	1.39***	-0.67***	0.49*	0.05	-0.03+	-0.79***	0.01
	(0.05)	(1.01)	(0.42)	(0.16)	(0.25)	(0.08)	(0.02)	(0.18)	(0.01)
Second Year	-0.33***	-9.99***	1.26***	-0.77***	0.15	0.17**	-0.01	-0.48***	-0.00
	(0.04)	(0.83)	(0.31)	(0.12)	(0.15)	(0.06)	(0.01)	(0.14)	(0.01)
Years 3-5	-0.26***	-8.54***	1.41***	-0.82***	0.24**	0.09+	0.01	-0.32***	0.00
	(0.02)	(0.37)	(0.18)	(0.07)	(0.09)	(0.05)	(0.01)	(0.07)	(0.00)
Teachers	20896	20896	20896	20896	20828	15463	20607	20893	20896
Observations	135175	135175	135175	135175	131093	60722	120239	134633	135175
Secondary Teachers									
First Year	-0.46***	-10.56***	-0.73*	-0.19+	-0.32	0.06	-0.09***	-1.04***	-0.01
	(0.05)	(0.64)	(0.32)	(0.10)	(0.24)	(0.05)	(0.02)	(0.12)	(0.01)
Second Year	-0.42***	-9.78***	-0.02	-0.22*	-0.08	0.07	-0.10***	-0.85***	-0.00
	(0.03)	(0.55)	(0.26)	(0.08)	(0.20)	(0.06)	(0.02)	(0.10)	(0.01)
Years 3-5	-0.27***	-6.76***	0.12	-0.18***	-0.12	0.00	-0.04***	-0.57***	-0.01+
	(0.02)	(0.33)	(0.15)	(0.05)	(0.14)	(0.05)	(0.01)	(0.06)	(0.00)
Teachers	15145	15145	15145	15145	15082	10353	14728	15145	15145
Observations	85454	85454	85454	85454	83379	37543	73969	85423	85454
Special Education Teachers									
First Year	-0.21***	-5.11***	-0.15	-0.34**	0.18	0.08*	-0.04**	-0.34**	0.01
	(0.03)	(0.48)	(0.30)	(0.11)	(0.20)	(0.04)	(0.02)	(0.12)	(0.01)
Second Year	-0.18***	-4.87***	-0.14	-0.01	0.33	0.18***	-0.04**	-0.31**	0.00
	(0.03)	(0.48)	(0.28)	(0.12)	(0.21)	(0.04)	(0.01)	(0.12)	(0.01)
Years 3-5	-0.11***	-3.19***	-0.25	-0.03	-0.04	0.09*	-0.03***	-0.08	0.00
	(0.02)	(0.28)	(0.18)	(0.07)	(0.13)	(0.04)	(0.01)	(0.08)	(0.00)
Teachers	7519	7519	7519	7519	7487	5813	7011	7516	7519
Observations	39622	39622	39622	39622	38505	19557	32266	39494	39622

Note. Standard errors clustered on school-by-year in parentheses. Includes teachers from 2007-8 through 2016-17. All models include year fixed effects.
+ p<.1, * p<.05 ** p<.01 *** p<.001

Table 7 – Teachers' VAMs as a Function of Composite Context Measures

	ELA VAM						Math VAM					
	All		Survey Sample				All		Survey Sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Instructional	-0.38***	-0.26***	-0.27***		-0.25***	-0.30***	-0.23***	-0.18***	-0.31***		-0.32***	-0.36***
Load	(0.02)	(0.02)	(0.04)		(0.04)	(0.06)	(0.02)	(0.02)	(0.04)		(0.04)	(0.08)
x Veteran	-0.02	0.05**	-0.14***		-0.14***	0.06	-0.04*	0.03	-0.05		-0.03	0.18*
	(0.02)	(0.02)	(0.04)		(0.04)	(0.06)	(0.02)	(0.02)	(0.04)		(0.04)	(0.08)
Homophily	0.01	0.02	0.01		0.01	0.17**	0.05***	0.02	-0.01		-0.00	0.13*
	(0.01)	(0.02)	(0.04)		(0.04)	(0.06)	(0.01)	(0.02)	(0.03)		(0.03)	(0.07)
x Veteran	0.02	0.03+	0.02		0.03	-0.10	-0.04*	0.02	0.00		0.00	-0.05
	(0.01)	(0.02)	(0.04)		(0.04)	(0.07)	(0.01)	(0.02)	(0.03)		(0.03)	(0.07)
Colleague	0.04*	-0.04*	0.08*		0.07+	-0.12*	0.04**	0.00	0.01		0.02	0.01
Qualifications	(0.02)	(0.02)	(0.04)		(0.04)	(0.06)	(0.01)	(0.02)	(0.03)		(0.03)	(0.05)
x Veteran	0.01	0.03+	-0.03		-0.03	0.05	0.01	-0.01	0.04		0.02	-0.05
	(0.02)	(0.02)	(0.04)		(0.04)	(0.06)	(0.01)	(0.02)	(0.03)		(0.03)	(0.05)
Professional				0.11**	0.06	-0.15*				0.03	-0.02	-0.03
Culture				(0.04)	(0.04)	(0.06)				(0.03)	(0.04)	(0.06)
x Veteran				-0.00	-0.01	0.15**				0.07*	0.06+	0.05
				(0.04)	(0.04)	(0.06)				(0.03)	(0.04)	(0.06)
Teacher FE		X				X		X				X
R-sq	0.13	0.56	0.13	0.04	0.13	0.64	0.06	0.62	0.09	0.03	0.10	0.68
Teachers	14838	11600	7408	7408	7408	4133	13678	10844	7223	7223	7223	4113
Observations	59609	56371	15226	15226	15226	11951	56805	53971	14919	14919	14919	11809

Note. All models are weighted by the number of measures observed and include indicators for school year, teachers' years of experience, and whether the teacher is elementary, secondary, or special education. Models without teacher fixed effects also include indicators for teacher race and gender. Standard errors clustered on school-by-year in parentheses.

+ p<.1, * p<.05 ** p<.01 *** p<.001

Table 8 – Teachers' Evaluation and Attendance Outcomes as a Function of Composite Context Measures

	Below Standard Evaluation[a]					Average EDST Rating						Attendance Rate					
	All		Survey Sample			All		Survey Sample				All		Survey Sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
Instructional	1.31***	1.07	1.24***		1.23***	-0.09***	-0.05*	-0.10***		-0.10***	-0.08**	-0.04	-0.17*	-0.02		0.00	-0.19*
Load	(0.06)	(0.92)	(0.08)		(0.08)	(0.01)	(0.02)	(0.02)		(0.02)	(0.03)	(0.04)	(0.07)	(0.04)		(0.04)	(0.08)
x Veteran	0.95	1.08	0.96		0.96	-0.03+	-0.02	-0.04*		-0.03+	-0.02	-0.19***	0.07	-0.21***		-0.20***	0.09
	(0.05)	(0.78)	(0.07)		(0.07)	(0.02)	(0.03)	(0.02)		(0.02)	(0.03)	(0.04)	(0.07)	(0.05)		(0.05)	(0.08)
Homophily	0.87*	0.91	0.94		0.92	0.01	0.02	0.00		0.01	0.02	0.13*	0.02	0.12*		0.14**	0.05
	(0.06)	(0.63)	(0.11)		(0.11)	(0.01)	(0.03)	(0.02)		(0.02)	(0.03)	(0.06)	(0.07)	(0.05)		(0.05)	(0.08)
x Veteran	1.05	1.17	1.10		1.10	0.05***	0.02	0.06***		0.06***	0.01	-0.11*	-0.02	-0.11*		-0.11*	-0.03
	(0.08)	(0.70)	(0.13)		(0.13)	(0.02)	(0.04)	(0.02)		(0.02)	(0.04)	(0.06)	(0.08)	(0.06)		(0.06)	(0.09)
Colleague	0.95	0.81	1.00		1.00	0.07***	0.03	0.06**		0.05**	-0.00	0.14***	0.08	0.18***		0.15**	0.20**
Qualifications	(0.05)	(0.44)	(0.10)		(0.10)	(0.02)	(0.02)	(0.02)		(0.02)	(0.03)	(0.04)	(0.06)	(0.05)		(0.05)	(0.07)
x Veteran	0.84**	1.11	0.72**		0.74**	0.04*	-0.00	0.06**		0.05*	0.03	0.01	-0.04	-0.04		-0.03	-0.18*
	(0.05)	(0.59)	(0.08)		(0.08)	(0.02)	(0.03)	(0.02)		(0.02)	(0.04)	(0.04)	(0.06)	(0.05)		(0.05)	(0.07)
Professional				0.97	0.99				0.07***	0.06***	0.07*				0.24***	0.21***	0.02
Culture				(0.10)	(0.11)				(0.02)	(0.02)	(0.04)				(0.05)	(0.05)	(0.07)
x Veteran				0.73**	0.74*				0.07***	0.07***	-0.04				0.03	0.03	0.10
				(0.08)	(0.09)				(0.02)	(0.02)	(0.04)				(0.05)	(0.05)	(0.08)
Teacher FE		X					X				X		X				X
R-sq						0.08	0.75	0.08	0.07	0.09	0.76	0.02	0.46	0.02	0.02	0.02	0.52
Teachers	34289	1047	17427	17427	17427	17298	4680	14613	14613	14613	3322	30397	26483	26763	26763	26763	21897
Observations	86645	4324	23289	23289	23289	22914	10296	18493	18493	18493	7202	129818	125904	79112	79112	79112	74246

Note. All models are weighted by the number of measures observed and include indicators for year, teachers' years of experience, and whether the teacher is elementary, secondary, or special education. Models without teacher fixed effects also include indicators for teacher race and gender. Standard errors clustered on school-by-year in parentheses.

[a] Logistic regressions. Coefficients are odds ratios.

+ p<.1, * p<.05 ** p<.01 *** p<.001

Table 9 – Average Marginal Probabilities of Teacher Mobility Outcomes as a Function of Composite Context Measures

	Stay				Switch School				Leave District			
	All	Survey Sample			All	Survey Sample			All	Survey Sample		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Instructional Load												
Novices	-1.09***	-1.32**		-1.18**	0.46*	0.22		0.12	0.62***	1.10***		1.06***
	(0.27)	(0.46)		(0.46)	(0.23)	(0.38)		(0.38)	(0.17)	(0.31)		(0.31)
Veterans	-1.49***	-1.12***		-1.04***	1.41***	1.01***		0.95***	0.08	0.11		0.09
	(0.09)	(0.14)		(0.14)	(0.08)	(0.11)		(0.11)	(0.05)	(0.09)		(0.09)
Homophily												
Novices	0.67*	0.60		0.68	0.19	-0.38		-0.43	-0.87***	-0.22		-0.25
	(0.30)	(0.58)		(0.57)	(0.24)	(0.42)		(0.42)	(0.20)	(0.42)		(0.42)
Veterans	0.94***	0.74***		0.79***	-0.58***	-0.54***		-0.56***	-0.36***	-0.20*		-0.23*
	(0.09)	(0.15)		(0.15)	(0.08)	(0.12)		(0.12)	(0.05)	(0.10)		(0.10)
Colleague Qualifications												
Novices	0.28	1.85***		1.52**	0.54*	-0.84*		-0.61	-0.82***	-1.02**		-0.90**
	(0.32)	(0.49)		(0.49)	(0.27)	(0.39)		(0.39)	(0.18)	(0.35)		(0.35)
Veterans	1.07***	1.38***		1.26***	-0.94***	-1.02***		-0.95***	-0.13**	-0.36***		-0.32***
	(0.13)	(0.16)		(0.15)	(0.12)	(0.13)		(0.13)	(0.05)	(0.08)		(0.08)
Professional Culture												
Novices			2.53***	2.13***			-1.38**	-1.26**			-1.15**	-0.86*
			(0.60)	(0.60)			(0.45)	(0.46)			(0.42)	(0.42)
Veterans			1.45***	1.25***			-1.03***	-0.84***			-0.42***	-0.41***
			(0.17)	(0.17)			(0.15)	(0.15)			(0.09)	(0.09)
Teachers	37950	26768	26768	26768	37950	26768	26768	26768	37950	26768	26768	26768
Observations	235177	79140	79140	79140	235177	79140	79140	79140	235177	79140	79140	79140

Note. All models are weighted by the number of measures observed and include indicators for year; teachers' race, gender, and years of experience; and whether the teacher is elementary, secondary, or special education. Models without teacher fixed effects also include indicators for teacher race and gender. Standard errors clustered on school-by-year in parentheses. Coefficients are average marginal effects for novices and veterans, respectively.

+ p<.1, * p<.05 ** p<.01 *** p<.001

Online Appendix A

Appendix Table A1 – Professional Culture Survey Questions Used

Response Options: Strongly Agree/Agree/[Neither Disagree Nor Agree]/Disagree/Strongly Disagree

1. What I learn in our school professional development meetings addresses my students' needs.
2. The professional development at this school is [appropriate] [differentiated] for my level of teaching experience.
3. The [current performance review] [Educator Development and Support process for teachers] helps me improve my teaching and learning.
4. I have sufficient autonomy to implement an instructional program that meets the needs of my students.
5. At this school, parents are partners with the school in decisions made about their children's education.
6. [At this school we trust one another.] [This school promotes trust and collegiality among staff.]*
7. I feel comfortable talking with the school leadership about issues and concerns.*

Response Options: Always/Often/Sometimes/Rarely/Never

8. In professional development this year: worked in grade-level or department-level teams to review and align grading practices. **
9. How often does school leadership visit your classroom to observe you teach?*
10. How often does school leadership provide useful feedback to you based on their observations?

Note. Text in brackets indicates alternate language used in some years, which we combine across years.

** All questions went to surveyed teachers only, with the exception of these questions in 2014-15, 2015-16, and 2016-17, which went to all surveyed staff.*

*** In 2016-17, the response options for these questions included: weekly/twice a month/monthly/a few times a year/hardly ever/never. We calculate the percent of respondents selecting these last three categories in our measure of agreement/frequency for these questions.*

Appendix Table A2 – New Teachers' Professional Culture Compared to Teachers in Years 6+

		Agree/Strongly Agree								Often/Very Often	
Elementary	Professional Culture Score (1)	PD Meets Students' Needs (2)	PD Appropriate for My Experience (3)	Performance Review Helps Improvement (4)	Teachers have Sufficient Autonomy (5)	Parents are Partners (6)	Trust Amongst Staff (7)	Comfortable Talking with Leaders (8)	There is Team Alignment (9)	Leaders Visit to Observe (10)	Leaders Provide Useful Feedback (11)
Experience Level (Reference Group = Teachers with 6 or more years experience)											
First Year	-0.07	-0.64	-1.17	1.07	-1.72^{+}	-2.40*	-2.97**	-1.96^{+}	-0.79	-0.04	-0.99
	(0.06)	(0.93)	(0.96)	(0.99)	(0.92)	(1.18)	(1.08)	(1.06)	(1.00)	(1.17)	(1.09)
Second Year	0.02	0.37	0.05	1.40	-1.00	0.46	-0.92	-0.29	0.58	1.93	0.06
	(0.06)	(0.94)	(0.89)	(0.90)	(1.03)	(1.11)	(1.04)	(0.99)	(1.03)	(1.25)	(1.06)
Years 3-5	-0.02	0.30	0.08	0.84	-1.42^{+}	-0.55	-0.96	-0.66	0.84	0.48	0.02
	(0.04)	(0.69)	(0.69)	(0.63)	(0.79)	(0.89)	(0.83)	(0.76)	(0.66)	(0.99)	(0.86)
Teachers	14577	14481	14489	14485	14477	14509	14187	14150	14470	14454	14454
Observations	46057	45060	45117	45004	45055	45404	42947	42507	44955	44930	44890
Secondary											
First Year	0.06	1.31	0.38	1.66*	0.41	-1.47	0.17	0.22	1.26	3.14**	0.90
	(0.05)	(0.94)	(0.93)	(0.82)	(0.94)	(1.03)	(1.10)	(1.02)	(0.82)	(1.11)	(0.88)
Second Year	0.04	1.49	1.18	1.63^{+}	0.35	-2.60*	0.07	0.07	-0.59	2.02*	1.11
	(0.06)	(1.13)	(0.99)	(0.86)	(0.95)	(1.03)	(1.25)	(1.13)	(0.91)	(1.02)	(0.98)
Years 3-5	0.11**	1.93**	1.53**	2.22***	2.00**	-1.32*	2.17**	2.03**	0.85	2.20**	1.33*
	(0.04)	(0.67)	(0.58)	(0.57)	(0.61)	(0.65)	(0.74)	(0.70)	(0.67)	(0.70)	(0.61)
Teachers	9501	9402	9403	9405	9396	9410	9347	9294	9403	9397	9385
Observations	26601	25982	25950	25981	25995	26110	24227	23705	25998	25904	25855
Special Education											
First Year	-0.03	-0.92	-1.17*	-0.13	-2.05**	-1.62*	-0.61	-0.82	0.89	1.45^{+}	-0.04
	(0.04)	(0.59)	(0.59)	(0.54)	(0.63)	(0.75)	(0.70)	(0.60)	(0.63)	(0.81)	(0.67)
Second Year	-0.05	-0.73	-1.20*	-0.39	-1.48*	-2.11**	-0.78	-0.30	0.52	0.57	-0.50
	(0.04)	(0.63)	(0.61)	(0.58)	(0.64)	(0.76)	(0.73)	(0.67)	(0.69)	(0.90)	(0.72)
Years 3-5	-0.01	-0.28	-0.50	0.23	-0.58	-1.15^{+}	0.25	0.17	1.01^{+}	-0.20	-0.42
	(0.03)	(0.49)	(0.48)	(0.44)	(0.46)	(0.61)	(0.56)	(0.49)	(0.56)	(0.66)	(0.56)
Teachers	5448	5410	5414	5412	5410	5413	5282	5259	5410	5406	5403
Observations	14793	14489	14485	14476	14482	14573	13561	13366	14473	14448	14437

Note. Standard errors clustered on school-by-year in parentheses. Includes teachers from 2007-8 through 2016-17. All models include year fixed effects.
$^{+}$ p<.1, * p<.05 ** p<.01 *** p<.001

Appendix Table A3 – New Teachers' VAM, Evaluation, and Attendance and Outcomes by Teacher Type

	ELA VAM		Math VAM		Below Standard Evaluation[a]		Average EDST Rating		Attendance Rate	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Elementary										
Instructional	-0.22***	-0.28***	-0.13***	-0.32***	1.50***	1.36+	-0.10*	-0.10*	-0.23**	-0.20*
Load	(0.02)	(0.07)	(0.02)	(0.08)	(0.13)	(0.24)	(0.04)	(0.05)	(0.08)	(0.10)
x Veteran	0.03	0.03	-0.00	0.14+	0.83*	0.87	0.01	0.00	0.16*	0.14
	(0.02)	(0.08)	(0.02)	(0.08)	(0.07)	(0.16)	(0.05)	(0.06)	(0.08)	(0.10)
Homophily	0.00	0.18*	-0.00	0.08	0.93	1.03	-0.08+	-0.05	0.05	0.22*
	(0.02)	(0.08)	(0.02)	(0.09)	(0.13)	(0.21)	(0.04)	(0.05)	(0.09)	(0.10)
x Veteran	0.03	-0.13	0.03	-0.03	1.01	1.04	0.14**	0.10+	0.02	-0.15
	(0.02)	(0.08)	(0.02)	(0.09)	(0.14)	(0.22)	(0.05)	(0.06)	(0.10)	(0.11)
Colleague	-0.04*	-0.08	-0.01	0.02	0.89	0.97	-0.02	-0.05	0.20*	0.23*
Qualifications	(0.02)	(0.06)	(0.02)	(0.07)	(0.10)	(0.16)	(0.04)	(0.04)	(0.09)	(0.11)
x Veteran	0.03	0.01	-0.00	-0.06	0.86	0.76	0.04	0.07	-0.17+	-0.20+
	(0.02)	(0.07)	(0.02)	(0.07)	(0.10)	(0.13)	(0.04)	(0.05)	(0.09)	(0.11)
Professional		-0.19**		-0.04		1.08		0.08		0.06
Culture		(0.07)		(0.07)		(0.22)		(0.07)		(0.11)
x Veteran		0.19**		0.04		0.65*		-0.03		0.07
		(0.07)		(0.07)		(0.13)		(0.07)		(0.11)
Teacher FE	X	X	X	X			X	X	X	X
Teachers	9026	3472	8911	3392	18146	9124	1994	1450	13904	11806
Observations	43623	10007	43162	9786	45318	11882	4360	3140	67574	41124
Secondary										
Instructional	-0.37***	-0.51***	-0.27***	-0.35	1.52**	1.13	-0.14*	-0.12	-0.08	-0.29
Load	(0.03)	(0.14)	(0.04)	(0.23)	(0.22)	(0.36)	(0.06)	(0.07)	(0.17)	(0.27)
x Veteran	0.10**	0.35*	0.04	0.16	0.86	1.11	0.01	-0.04	-0.24	-0.04
	(0.04)	(0.15)	(0.05)	(0.23)	(0.13)	(0.37)	(0.07)	(0.09)	(0.17)	(0.27)
Homophily	0.06+	0.10	0.07*	0.30*	0.76*	0.49**	0.05	0.10	-0.10	-0.11
	(0.04)	(0.09)	(0.03)	(0.13)	(0.09)	(0.12)	(0.06)	(0.07)	(0.11)	(0.16)
x Veteran	0.04	0.02	0.02	-0.03	1.23	2.10**	-0.06	-0.16+	-0.09	-0.06
	(0.03)	(0.09)	(0.03)	(0.14)	(0.15)	(0.53)	(0.07)	(0.09)	(0.12)	(0.18)
Colleague	-0.02	-0.28*	0.03	0.04	0.95	1.00	0.07	0.02	0.06	0.23*
Qualifications	(0.03)	(0.13)	(0.03)	(0.09)	(0.08)	(0.17)	(0.04)	(0.05)	(0.08)	(0.11)
x Veteran	0.02	0.22+	0.01	-0.01	0.87	0.72+	-0.02	0.01	-0.02	-0.24+
	(0.03)	(0.13)	(0.03)	(0.10)	(0.08)	(0.13)	(0.05)	(0.07)	(0.09)	(0.13)
Professional		-0.02		0.13		0.74+		0.06		0.10
Culture		(0.10)		(0.15)		(0.13)		(0.06)		(0.14)
x Veteran		0.11		-0.01		1.02		-0.02		0.01
		(0.09)		(0.14)		(0.21)		(0.07)		(0.15)
Teacher FE	X	X	X	X			X	X	X	X
Teachers	2695	634	2066	698	12278	5590	1520	1012	9072	7120
Observations	12176	1775	10337	1873	29223	7377	3330	2173	42094	23403
Special Education										
Instructional					0.08	0.08	0.01	-0.04	0.89	0.92
Load					(0.08)	(0.08)	(0.03)	(0.04)	(0.09)	(0.08)
x Veteran					0.10	0.08	-0.01	-0.02	1.19	1.10
					(0.10)	(0.12)	(0.04)	(0.07)	(0.13)	(0.13)
Homophily					0.05	0.22	0.08+	0.04	1.04	1.01
					(0.11)	(0.18)	(0.05)	(0.05)	(0.16)	(0.14)
x Veteran					-0.09	-0.15	-0.00	0.06	1.05	1.15
					(0.12)	(0.19)	(0.06)	(0.08)	(0.19)	(0.20)
Colleague					0.00	0.01	0.02	0.03	1.11	1.16
Qualifications					(0.11)	(0.19)	(0.04)	(0.05)	(0.14)	(0.17)
x Veteran					-0.16	-0.23	-0.03	0.00	0.98	0.93
					(0.14)	(0.22)	(0.06)	(0.08)	(0.14)	(0.16)
Professional						0.06		0.06		0.90
Culture						(0.19)		(0.05)		(0.13)
x Veteran						-0.26		-0.09		1.42+
						(0.24)		(0.08)		(0.28)
Teacher FE							X	X	X	X
Teachers					5220	2736	1069	797	3766	2969
Observations					12104	3800	2354	1721	15482	9176

Note. All models are weighted by the number of measures observed and include indicators for year and teachers' years of experience. Models without teacher fixed effects also include indicators for teacher race and gender. Standard errors clustered on school-by-year in parentheses.

[a] Logistic regressions. Coefficients are odds ratios.

+ p<.1, * p<.05 ** p<.01 *** p<.001

Appendix Table A4 –Teachers' Mobility Outcomes as a Function of Composite Context Measures, by Teacher Type

		Stay		Switch School		Leave District	
		(1)	(2)	(3)	(4)	(5)	(6)
Elementary							
Instructional Load:	Novices	-2.70***	-1.22	2.03***	1.68*	0.67*	-0.46
		(0.46)	(0.94)	(0.39)	(0.69)	(0.28)	(0.70)
	Veterans	-0.98***	-0.42**	0.90***	0.36**	0.08	0.06
		(0.10)	(0.16)	(0.08)	(0.12)	(0.06)	(0.11)
Homophily:	Novices	0.88*	1.65+	0.30	-0.40	-1.18***	-1.25+
		(0.45)	(0.92)	(0.41)	(0.75)	(0.27)	(0.65)
	Veterans	0.58***	0.67***	-0.19*	-0.35*	-0.39***	-0.33**
		(0.11)	(0.17)	(0.09)	(0.14)	(0.07)	(0.12)
Colleague Qualifications:	Novices	-0.93*	0.27	1.53***	0.67	-0.60*	-0.94
		(0.44)	(0.82)	(0.41)	(0.66)	(0.25)	(0.64)
	Veterans	0.46***	0.81***	-0.42***	-0.58***	-0.05	-0.23*
		(0.11)	(0.15)	(0.10)	(0.12)	(0.06)	(0.09)
Professional Culture:	Novices		3.88***		-2.73***		-1.15+
			(0.89)		(0.76)		(0.66)
	Veterans		1.27***		-1.02***		-0.25*
			(0.16)		(0.14)		(0.10)
Teachers		19797	13951	19797	13951	19797	13951
Observations		125557	43272	125557	43272	125557	43272
Secondary							
Instructional Load:	Novices	-1.89**	-0.94	2.25***	0.30	-0.36	0.64
		(0.59)	(1.34)	(0.48)	(1.14)	(0.40)	(0.96)
	Veterans	-2.69***	-2.33***	2.53***	2.18***	0.16	0.15
		(0.23)	(0.33)	(0.21)	(0.28)	(0.11)	(0.21)
Homophily:	Novices	2.12***	1.43	-0.62	-0.13	-1.50***	-1.30
		(0.50)	(1.08)	(0.38)	(0.78)	(0.33)	(0.82)
	Veterans	1.71***	1.37***	-1.26***	-1.18***	-0.45***	-0.19
		(0.22)	(0.36)	(0.18)	(0.29)	(0.12)	(0.22)
Colleague Qualifications:	Novices	1.05*	2.78**	-0.21	-2.00**	-0.84**	-0.78
		(0.50)	(0.86)	(0.43)	(0.71)	(0.27)	(0.53)
	Veterans	1.73***	2.06***	-1.49***	-1.57***	-0.24**	-0.49**
		(0.24)	(0.30)	(0.22)	(0.26)	(0.09)	(0.17)
Professional Culture:	Novices		2.66*		-2.07*		-0.60
			(1.19)		(0.82)		(0.83)
	Veterans		1.25***		-0.41		-0.85***
			(0.37)		(0.32)		(0.20)
Teachers		14454	9213	14454	9213	14454	9213
Observations		81088	25508	81088	25508	81088	25508
Special Education							
Instructional Load:	Novices	-1.34**	-1.15	0.38	-0.12	0.96**	1.27**
		(0.45)	(0.74)	(0.35)	(0.61)	(0.30)	(0.49)
	Veterans	-0.88***	-1.04**	0.64***	0.57*	0.24+	0.48*
		(0.23)	(0.34)	(0.19)	(0.29)	(0.13)	(0.21)
Homophily:	Novices	0.73	0.89	-0.62	-1.40*	-0.10	0.50
		(0.54)	(0.93)	(0.40)	(0.67)	(0.40)	(0.71)
	Veterans	0.57*	0.05	-0.14	0.34	-0.43**	-0.39
		(0.26)	(0.45)	(0.21)	(0.36)	(0.16)	(0.28)
Colleague Qualifications:	Novices	1.09*	1.29	0.10	0.10	-1.19***	-1.39*
		(0.50)	(0.87)	(0.37)	(0.64)	(0.36)	(0.69)
	Veterans	0.89**	0.67	-0.79***	-0.44	-0.10	-0.23
		(0.28)	(0.42)	(0.23)	(0.32)	(0.14)	(0.25)
Professional Culture:	Novices		0.91		-0.02		-0.90
			(0.94)		(0.73)		(0.70)
	Veterans		0.94*		-0.63+		-0.32
			(0.46)		(0.37)		(0.28)
Teachers		5990	4145	5990	4145	5990	4145
Observations		28532	10360	28532	10360	28532	10360

Note. All models are weighted by the number of measures observed and include indicators for year and for teachers' race, gender, and years of experience. Standard errors clustered on school-by-year in parentheses. Coefficients are average marginal effects for novices and veterans, respectively.

+ p<.1, * p<.05 ** p<.01 *** p<.001

Online Appendix B – VAM Estimation

Linking Students to Teachers

District report card files link students to teachers in each subject/trimester/year at the elementary level and in each class period/subject/semester/year at the secondary level. At the elementary level, students are linked for VAM estimation purposes to any teacher listed as the teacher of record for the appropriate subject (math or ELA) including corresponding English Language Development (ELD) courses. At the secondary level, teachers of math and ELA courses are linked to students only if the course is not more advanced than the content on the end-of-year test taken by the student. A student can be linked to a teacher in multiple instances if the student takes multiple same-subject courses in different class periods (e.g., a grade-level course and an intervention course). To ensure that teachers have adequate instructional time with students to justify the estimation of a VAM, we require that teachers be linked to a student at least in the first trimester or semester, and at the elementary level in at least one additional trimester. Each teacher-student link is assigned a weight corresponding to the share of the year for which the student and teacher were linked in the administrative data (i.e., 0.33 per trimester or 0.5 per semester; Hock & Isenberg, 2017).

We do not use classrooms (school-teacher-year combinations at the elementary level and school-teacher-class period-year combinations at the secondary level) where 60 percent or more of students are designated as eligible for special education services. Additionally, teachers must be connected to at least eight unique tested students with complete (e.g., prior achievement) data, and must be connected to at least eight full-year-equivalent student records (e.g., a teacher linked to 16 tested students each for a single semester could have a VAM estimated, but a teacher linked to eight students each for a single semester would not).

VAM Model

We estimate model (1) for each year/subject/level combination:

$$Ach_{ijst} = \beta_1 Ach^{math}_{ijst-1} + \beta_2 Ach^{ela}_{ijst-1} + \boldsymbol{X}_{ijst}\boldsymbol{\theta} + \boldsymbol{T}_{jt}\boldsymbol{\Omega} + \varepsilon_{ijst} \quad (1)$$

where Ach is either math or ELA achievement, standardized within test and year, for student i with teacher j in school s in year t. We control for students' achievement in the prior year in both math and ELA. $\boldsymbol{X}$ is a vector of student demographic characteristics, including indicators of student race, gender, free- or reduced-price lunch eligibility, special education status, English learner status, and grade level. Teachers' VAMs are estimated by the coefficients on a set of teacher fixed effects ($\boldsymbol{T}$). ε is an error term. Each observation is weighted as described above. We estimate heteroskedasticity-robust standard errors. Estimates are then shrunken using Empirical Bayes (EB) shrinkage (Herrmann, Walsh, Isenberg, & Resch, 2013; Koedel et al., 2015). VAMs estimated with heteroskedasticity-robust standard errors and VAMs estimated with standard errors clustered on students correlate at $r. > .99$ after shrinkage across years for both subjects at both levels.

Missing Achievement Data in 2013-14

The estimation of these VAMs is complicated by the fact that after the 2012-13 school year California suspended the use of its Standardized Testing and Reporting (STAR) accountability system and replaced it with the California Assessment of Student Performance and Progress (CAASPP), a new set of math and ELA tests aligned to the Common Core State Standards. This transition included one school year – 2013-14 – in which the new tests were piloted but no results were reported. To estimate VAMs in the 2014-15 school year we therefore use as controls students' achievement two years prior (2012-13). Whether this is sufficient to estimate unbiased VAMs is not obvious, particularly given that the pre-test and post-test derive

from separate testing regimes, though the results presented here are generally similar if estimated separately for CST and CAASPP or excluding the 2014-15 school year. Additionally, when it is possible (i.e., in other years) to estimate VAMs using either achievement in *t*-1 or *t*-2, we find that those estimates have high correlations, ranging from $r = .79$ (elementary ELA) to $r = .94$ (secondary math) across all years. The stability of our results to the inclusion of multiple testing regimes is consistent with prior research finding that VAMs are fairly stable across transitions between testing regimes, particularly in math (Backes et al., 2018).

Alternate Specifications

Controlling for an additional lag of same-subject achievement can enhance model performance (Ehlert, Koedel, Parsons, & Podgursky, 2014). This substantially reduces the number of teachers for whom we can estimate VAMs, but VAMs estimated in this way correlate at $r > .92$ with those from our primary specification in both subjects and levels.

An additional concern is that model (1) may not fully control for sorting of teachers to students or for effects of students' peers. We therefore estimate two additional specifications that control for three characteristics of other students (in the classroom or school, respectively) used by Isenberg et al. (2016), including average peer prior achievement in the same subject, the standard deviation of peer prior achievement in the same subject, and the share of peers who are eligible for free- or reduced-price lunch.

In our primary specification we estimate single-year VAMs, allowing each teacher's VAM to vary arbitrarily from year to year, but often leaving little within-teacher variation in school or classroom demographics. We therefore check for peer effects by pooling two years of data, rather than one, and including a school year dummy variable in the model. These VAMs estimated with and without peer effects correlate highly (approximately $r = .95$), consistent with prior work (e.g., Johnson, Lipscomb, & Gill, 2015).

References

Backes, B., Cowan, J., Goldhaber, D., Koedel, C., Miller, L. C., & Xu, Z. (2018). The Common Core Conundrum: To What Extent Should We Worry That Changes to Assessments Will Affect Test-Based Measures of Teacher Performance? *Economics of Education Review*, *62*, 48–65.

Ehlert, M., Koedel, C., Parsons, E., & Podgursky, M. (2016). Selecting Growth Measures for Use in School Evaluation Systems Should Proportionality Matter? *Educational Policy*, *30*(3), 465–500.

Herrmann, M., Walsh, E., & Isenberg, E. (2016). Shrinkage of Value-Added Estimates and Characteristics of Students with Hard-to-Predict Achievement Levels. *Statistics and Public Policy*, *3*(1), 1–10.

Hock, H., & Isenberg, E. (2017). Methods for Accounting for Co-Teaching in Value-Added Models. *Statistics and Public Policy*, *4*(1), 1–11.

Isenberg, E., Max, J., Gleason, P., Johnson, M., Deutsch, J., & Hansen, M. (2016). *Do Low-Income Students Have Equal Access to Effective Teachers? Evidence from 26 Districts* (No. NCEE 2017-4007). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Koedel, C., Mihaly, K., & Rockoff, J. E. (2015). Value-added modeling: A review. *Economics of Education Review*, *47*, 180–195.

Johnson, M. T., Lipscomb, S., & Gill, B. (2015). Sensitivity of Teacher Value-Added Estimates to Student and Peer Control Variables. *Journal of Research on Educational Effectiveness*, *8*(1), 60–83.