**** PRELIMINARY DRAFT****

**** PLEASE DO NOT CITE WITHOUT AUTHORS' PERMISSION ****

Teacher Quality Gaps by Disability and Socioeconomic Status: Evidence from Los Angeles

By Ijun Lai, W. Jesse Wood, Scott A. Imberman, Katharine O. Strunk, and Nathan D. Jones*

Abstract

While most students with disabilities (SWDs) receive instruction from general education teachers, little empirical work has investigated whether these students have suitable access to high-quality teachers. We explore the differences in teacher quality experienced by SWDs and students without disabilities (non-SWDs) in the Los Angeles Unified School District, examining how access varies within schools as well as across school-level disadvantage rates. We leverage several different indicators of teacher effectiveness for general education teachers who instruct both SWDs and non-SWDs. We find that SWDs are significantly more likely to have teachers with lower math value-added (-0.024 standard deviations) than their non-SWD peers and we find emerging gaps in teacher evaluation scores and exposure to novice teachers. In general, these gaps do not vary by school-level disadvantage.

*Lai: Mathematica, 505 14th Street, Suite 800, Oakland, CA 94612-1475 (email: ilai@mathematica-mpr.com); Wood: Department of Economics, Michigan State University, 486 W. Circle Drive, 110 Marshall-Adams Hall, East Lansing, MI 48824 (email: woodwil2@msu.edu); Imberman: Department of Economics, Michigan State University, 486 W. Circle Drive, 25D Marshall-Adams Hall, East Lansing, MI 48824 (email: imberman@msu.edu); Strunk: Department of Educational Administration, Michigan State University, 620 Farm Ln, 238 Erickson, East Lansing, MI 48824 (email: kstrunk@msu.edu); Jones: Wheelock College of Education & Human Development, Boston University, 2 Silber Way, SED 231, Boston, MA 02215 (email: ndjones@bu.edu)

Introduction

Since the 1975 passage of the Individuals with Disabilities Education Act (IDEA), federal law has dictated that students with disabilities (SWDs) be provided with a "free and appropriate public education in the least restrictive environment." In 2017, the U.S. Supreme Court established an even higher standard for special education with its decision in *Endrew F. v. Douglas County School District RE-1*. The case stressed the need to ensure equitable outcomes for SWDs, who continue to lag behind their nondisabled peers in math and reading achievement (e.g., Chudowsky, Chudowsky, & Keber, 2009; Schulte & Stevens, 2015; Schulte, Elliott, Tindal, & Nese, 2016). Today, roughly 6.4 million public school students in the U.S. receive special education services annually, and the majority receive their primary instruction in general education (U.S. Department of Education, 2019). Yet, to date, little empirical work has examined teacher quality for SWDs in the general education classroom context, raising critical questions about teacher quality gaps (TQGs) between students with disabilities in general education classrooms and their peers.

The teacher quality literature has repeatedly documented the unequal distribution of teachers—both within and across schools. Research shows that within schools, students sorting to teachers of varying quality often depends on their academic and behavioral histories (e.g. Kalogrides & Loeb, 2013; Lankford, Loeb, & Wyckoff, 2002). Furthermore, Goldhaber, Lavery, and Theobald (2015) and Goldhaber, Quince, and Theobald (2018) establish that substantial gaps exist in teacher quality across the socioeconomic distribution; low-income students consistently have less access to highly qualified teachers, likely because disadvantaged schools have more difficulty attracting and retaining high quality teachers (e.g., Boyd, Lankford, Loeb, & Wyckoff, 2005). Unfortunately, we do not have a parallel literature on TQGs experienced by students with disabilities.

It is unclear, based on existing research, whether to expect between- or within-school TQGs between SWDs in general education classrooms and students without disabilities. Regarding within-school gaps, there is some evidence to suggest that schools distribute SWDs to teachers in non-random ways. One recent study using North Carolina data (Gilmour & Henry, 2020) found that SWDs were more likely to have classmates with lower prior academic performance than their non-SWD peers. However, it is not clear whether this sorting results in differential access to high-quality teachers. On the one hand, it is possible that SWDs could be seen as more “difficult” students to teach and consequently, more likely to be assigned to lower quality teachers (e.g. Clotfelter, Ladd, & Vigdor, 2016). At the same time, because in many states and districts there is a higher level of accountability or attention paid to the opportunities given to SWDs (see, for example, Swaak, 2020), it could be the case that districts proactively assign SWDs to particularly effective teachers. Indeed, using a variety of teacher quality measures, Gilmour and Henry (2018) find little evidence of TQGs for SWDs overall in North Carolina. The gaps they did observe were due to within-school differences in teachers’ prior academic achievement. They also noted that a focus on SWDs overall masked some heterogeneity across disability subcategories.

To ensure equitable outcomes for SWDs, it is also important to identify between-school variation in teacher quality. The question of how TQGs vary by school-level disadvantage was not addressed in Gilmour and Henry (2018). We hypothesize that teacher quality gaps in high poverty schools may be felt even more acutely by students with disabilities. Gilmour and Wehby (2019) demonstrate that the likelihood of teacher turnover increases with the number of SWDs in the classroom. Given the higher concentration of SWDs in higher poverty schools, this finding suggests that SWDs in disadvantaged schools may be even less likely to access high-quality teachers than both non-SWDs in high poverty schools and SWDs in low poverty schools. However, to date, no studies have directly examined whether, within and across schools of varying income levels, TQGs exist across students with and without disabilities.

To promote more equitable outcomes for SWDs, we join these two strands of research and ask the following research questions in the Los Angeles Unified School District (LAUSD) context:

1) *Does teacher quality vary across schools with differing degrees of disadvantage?*
2) *Are there SWD vs non-SWD Teacher Quality Gaps?*
3) *Do SWD vs non-SWD gaps vary by school-level disadvantage?*
4) *Do teacher quality gaps across school-level disadvantages vary by specific disability type?*

This paper makes two primary contributions to the literature on TQGs among SWDs. First, we include multiple teacher quality indicators including value-added measures (VAMs), teachers' observation-based performance ratings, hiring scores, and teacher experience (novice status). This allows us to examine whether TQGs exist across a large, urban context (LAUSD) using an expanded range of quality measures. Second, we examine whether the small overall quality gaps found by Gilmour and Henry (2018) might mask school-level variation related to students' socioeconomic status. If higher-quality teachers are sorting into schools with fewer low socioeconomic students, we might find that TQGs are exacerbated across schools within the same district. Therefore, it is important to understand how TQGs might differ across schools with different degrees of student disadvantage rather than just overall differences within a district or state.

Data

Context

The Los Angeles Unified School District (LAUSD) is the second largest district in the country with approximately 570,000 K-12 students and 24,000 K-12 teachers in 2018. LAUSD administrative data provide detailed information on student and teacher characteristics and allow for student-teacher matches. The district is also economically and racially diverse, providing variation in teacher quality, school level wealth and achievement, disability status, and student characteristics.

The LAUSD classroom assignment process allows researchers a unique opportunity to study whether SWD TQG exists, because while principals assign teachers to classrooms, teachers have opportunities to request assignments. Teachers' assignment to classes are initially based on credentials. At the elementary level, teachers may submit requests for track and grade level positions. Teachers may be assigned to their preferred classes based on district seniority, though principals may dispute specific assignments if they believe that the assignment is not in the best interest of the school. At the secondary level, teachers submit requests for department selection, and principals consult with department heads to assign teachers to specific classes. See Appendix A for more details about teacher assignments.

Elementary classroom rosters are created at the end of the school year by grade-level teams. After the start of school, grade-level teams can call meetings to ensure that students are equitably distributed across classrooms. If they find the distribution unequal, they can recommend changes to the principal. Student-teacher pairings for elementary school classes and core classes in middle school are "fairly randomized" and placement adjustments are mostly around balancing class sizes (LAUSD, personal communication, December 12, 2019).

In LAUSD, approximately 60% of students with mild or moderate disabilities spend most of their days in general education classrooms (Swaak, 2020). The decision of whether a student should be included in general education is made on a case-by-case basis by a team of school personnel and outside professionals, in collaboration with parents. This decision is driven by the child's individualized needs. Then, among SWDs who are educated in the general education classroom, placement procedures into specific teachers' classrooms are decided by school personnel.

Sample

This study uses student- and teacher-level matched administrative data from SY2014-2015 through SY2017-2018. Data are provided by LAUSD's Office of Data and Accountability and the Division of Human Resources. Our sample includes all kindergarten through 8th grade students attending mainstream public schools during these years.[1] The data are at the student-year level and include demographic information such as disability status (detailed below), race/ethnicity, gender, free- or reduced-price lunch (FRL), and English Language Learner (ELL) status, as well as state standardized math and English Language Arts (ELA) test scores for students in grades 3-8. We normalize each subject's test scores to have a mean of zero and standard deviation of one for each grade-year combination. We include both math and ELA scores because of evidence that SWDs may have different challenges in each subject (e.g., Child et al., 2019; Fuchs et al., 2016). The data also contain teachers' demographic information, educational background, and contract status (i.e. pre-tenure and permanent). Additionally, the teacher files include teachers' final evaluation scores as well as observation subcomponent scores and, for teachers hired since 2013-14, their hiring scores from the district's teacher screening system.

Students are linked to teachers through a transcript file, which provides details on students' classroom placements and teacher for each class period.[2] The final dataset is restricted to students who are linked to at least one teacher. Our study focuses on SWDs who are taught by general education teachers since the majority of SWDs are in general education classrooms for most of their school days. Our overall sample consists of 1,175,666 student-year observations, or 13,107 unique teachers in 619 schools.

Variables of interest

Disability Status

We created indicator variables for four disability subcategories—autism, specific learning disability, speech/language impairment, and other. The categories reflect the disability groups that have high incidence rates in LAUSD and represent students with a range of needs. These categories are not mutually exclusive since students may have multiple disabilities.[3]

School Characteristics

Since previous literature has shown that teacher quality can vary across schools with different characteristics, we generate school-level characteristics at the year-level and then average across the four years in our panel (SY2014-2015 through SY2017-2018). Our main analysis focuses on school free- or reduced-price lunch (FRL) status.[4] We split schools into three categories based on their four-year FRL average: less than 70% FRL, 70%-<95%, and 95%-100%. We chose these FRL categories based on a combination of how previous literature has examined the FRL distribution and the distribution of FRL students within LAUSD, which skews towards high rates of poverty.[5] Grouping schools in this manner allows us to compare students with and without disabilities at schools with similar demographics, while also observing how these differences may vary across schools with different student characteristics.

Teacher Characteristics

The literature suggests that teacher input variables, such as teachers' educational histories and credentials, are poorly correlated with teacher effectiveness (e.g., Angrist & Guryan, 2008; Chingos & Peterson, 2011; Kane, Rockoff, & Staiger, 2008). Much of the more recent literature has advocated for using teacher output measures as indicators of teacher quality, such as VAMs and teacher evaluation scores (e.g., Aaronson, Barrow, & Sander, 2007; Rivkin et al., 2005). Additional research has shown that exposure to early career teachers has negative impacts on student performance (Clotfelter, Ladd, & Vigdor, 2007; Rice, 2010; Ladd & Sorensen, 2017; Staiger & Rockoff, 2010). Consequently, our main analysis focuses on four aspects of teacher quality: value-added measures (VAMs) of teachers' contributions to student achievement gains, teachers' ratings on their observation-based performance evaluations, teachers' initial hiring scores, and new teacher status (in first two years).

Teacher Value-Added Measures (VAMs)

We calculate VAMs for teachers teaching fourth through eighth grade. Following Chetty, Friedman, and Rockoff (2014), we use a multi-step calculation to create our value-added estimator.[6] First, to create residualized test scores, we regress scores on student, classroom, and grade demographics, including student-, classroom- and grade-level averages for race, gender, free/reduced lunch status, English Language Learner status, disability status, testing accommodation,[7] and previous test scores (cubed) in both ELA and math. Next, we average residualized test scores across all students for each teacher j in year t. We then calculate forecasting coefficients, which minimizes the mean squared error of the test-score predictions. Finally, we use data for teacher j in years outside of t to predict the value-added for teacher j in year t. See Chetty et al., 2014 and Appendix B for a more detailed description.

Teacher Evaluation Scores

Beginning in SY2014-2015, LAUSD teachers have been evaluated via classroom observations. All teachers new to a school are evaluated during their first two years. After the first two years, teachers are evaluated at least every other year, but some veteran teachers may extend the time between evaluations to up to five years.[8] In evaluation years, teachers are observed 1 or 2 times and receive scores on between 7 and 15 subcomponents from the Teaching and Learning Framework (depending on the year), as well as an overall evaluation score. Three subcomponents are required for all teachers across all years, while others are selected by teachers before the observation period. Since observation components varied by academic year and across teachers, we take the average score across all subcomponents and standardize by year.[9] In our sample, about 25-30% of our teachers are evaluated every year. Teachers who do not pass their evaluation are re-evaluated in the following year. We use teacher's evaluation scores from the prior year (or, for those who were not evaluated in the prior year, the most recent evaluation score) to create our teacher evaluation score measure.

Hiring Scores

In SY 2014-2015, LAUSD introduced a new teacher screening system. For teachers hired/re-hired 2014-2015, we have composite hiring scores based on application information (e.g., licensure, grade point averages), professional references, writing samples, interviews, and sample lesson demonstrations. Details can be found in Appendix C and Bruno and Strunk (2019). All hiring scores are standardized by year.

Experience

The current literature suggests that, on average, early career teachers rapidly improve their effectiveness over their first few years (e.g., Papay & Kraft, 2015; Kane et al., 2008; Rivkin et al., 2005), but that new teachers are generally lower quality than more experienced teachers. Consequently, we examine students' exposure to novice teachers, which we define as having two or fewer years of experience.[10]

Teacher Characteristics

Table 1 provides average general education teacher characteristics across FRL school bins and teacher quality measures.[11] Panels A through D highlight that each teacher quality measure is coming from a different teacher subsample, with "Novice" teachers as the most inclusive sample. Specifically, only teachers who taught grades four through eight will have VAM scores, only teachers who have been evaluated will have evaluation scores, and only teachers who have been hired (or re-hired) since SY2014-2015 will have hiring scores. To give a sense of how the sample varies across teacher quality measures, we also include descriptive statistics about the share of teachers with valid measures and the average score for these measures. For example, Panel A documents that about 32% of math teachers in our VAM sample have a valid teacher evaluation score and that the average z-scored evaluation score for this sample is 0.16.

There are many patterns that are consistent across all subsamples. Most notably, students at lower FRL (i.e., higher income) schools tend to have more white and female teachers than students at higher FRL schools. Additionally, on average, students in lower FRL schools tend to be exposed to teachers with higher teacher evaluation scores and hiring scores than students attending higher FRL schools.

Methods

We examine average teacher characteristics and whether these differ by student disability status. Our analytic approach is similar to that used in previous TQG literature (e.g., Clotfelter et al., 2005; Goldhaber et al., 2015). We use a simple bivariate regression of the following form:

$$Y_{ijs} = \beta_0 + \beta_1 Disability_{ijs} + \varepsilon_{ijs} \quad (1)$$

where Y_{ijs} represents the teacher quality measure of interest (i.e. <= 2 years of experience) for student *i* matched with teacher *j* at school *s*. We run the model separately by each disadvantage bin (FRL < 70%, 70-<95%, and 95+%). For our main results, *Disability*$_{ijs}$ is our SWD indicator variable. For our subgroup analysis, *Disability* represents one of three disability subgroups (Specific Learning Disability, Autism, or Speech/Language Impairment)[12] and zeros are given for non-SWDs. Standard errors are clustered at the school level. The bivariate regression allows us to calculate exposure rates to students with disabilities for teachers across different quality measures, as well as the exposure rates for their non-disabled peers, and to test if this difference (captured in β_1) is statistically significant. We also run a school fixed effects model to test the stability of our coefficients when focusing on within-school variation. Since results across the models are similar, we conclude that much of the TQGs are driven by within-school sorting, and only report coefficients from the unadjusted model (equation 1). We include variance decomposition information in Table 3 and Appendix Table 2 to examine the proportion of variation in TQGs that occurs within and across schools.

In addition to TQGs within FRL bins, we are interested in whether TQGs differ across bins (referred to as "disadvantage gaps" from this point forward). Specifically, we evaluate whether any of the TQGs are significantly different from the TQG in schools with less than 70% FRL students. To do this, we pool observations across two bins (with the most advantaged school as the reference bin) and estimate the following equation:

$$Y_{ijs} = \beta_0 + \beta_1 Disability_{ijs} + \beta_2 Adv\ Sch_{ijs} + \beta_3 (Disability_{ijs} * Adv\ Sch_{ijs}) + \varepsilon_{ijs} \quad (2)$$

where, again, Y_{ijs} represents the teacher quality measure of interest (i.e. <= 2 years of experience) for student *I* in teacher *j* at school *s* and *Disability*$_{ijs}$ is an indicator variable for students with disabilities. *Adv Sch*$_{ijs}$ is an indicator variable for the most advantaged school bin (FRL <70%). The *Disability*$_{ijs}$**Adv Sch*$_{ijs}$ interaction measures the teacher quality gap differences between the two school disadvantage bins and tests whether this difference is statistically significant. In the interest of space, we only display the p-value associated with *Disability*$_{ijs}$**Adv Sch*$_{ijs}$ in our tables.

Results

Research Question 1: Does teacher quality in LAUSD vary across schools with differing degrees of disadvantage?

Table 2 presents the mean and standard deviation for each general education teacher characteristic, split by subject, within our three FRL bins. Consistent with previous studies (e.g., Clotfelter et al, 2007; Goldhaber et al, 2018), we find some evidence of increasing exposure to lower quality teachers as we move down the column from most to least advantaged schools. For example, the average math teacher evaluation score for students attending the most advantaged schools in our sample (<70% FRL) was 0.361 (measured in standard deviation units), while the average score at the least advantaged school (>=95% FRL) was 0.161. For students in the middle FRL bin (70- <95% FRL), we find significant disadvantage gaps for both math and ELA VAMs. These students tend to be exposed to lower VAM teachers than students attending the lowest FRL schools, though we find no significant differences between students in the lowest and highest FRL bins. Additionally, students attending the middle and highest FRL bins tends to be exposed to teachers with lower teacher evaluation scores than their peers attending the lowest FRL schools. Aside from ELA teachers in the highest FRL bin, which have significantly lower hiring scores than their peers in the lowest FRL bin, hiring scores seem to be evenly distributed across subject and FRL bins. In contrast to studies in other contexts (e.g. Boyd et al, 2008), novice teachers are relatively equitably distributed across LAUSD schools regardless of school-level disadvantage.

Research Question 2: Are there SWD vs non-SWD TQGs? Do these gaps vary by school-level disadvantage?

Table 3 presents overall SWD vs non-SWD TQGs. We begin by examining the average teacher quality for SWD in general education classrooms and then the average teacher quality for their non-SWD peers across each of our teacher quality measures. Rows three and four present the quality gap and the corresponding standard error. Panel A presents our results for math teachers while panel B displays our ELA results. We find that, relative to students without disabilities, SWDs are assigned to lower quality math teachers in terms of VAMs and teacher evaluation scores. On average, SWDs in general education classrooms have math teachers with 0.024 standard deviation lower VAMs and 0.028 lower standardized teacher evaluation scores than their non-SWD peers. There are no significant gaps in experience (novice status) or hiring scores. For ELA teachers, SWDs tend to be assigned to teachers with 0.042 standard deviations (SD) lower evaluation scores and are more likely assigned to a novice teacher (0.6 pp).

The last two rows of each panel present estimates from a model adding a school fixed effect, which allows us to examine how much of the TQG variance is due to within- or between-school factors. For math, the VAM, teacher evaluation, and novice measures, approximately 2/3 of the gaps are driven by within-school differences, suggesting that the gaps are mostly a function of within-school distribution of teachers to SWDs and non-SWDs, rather that teacher sorting across schools. For hiring scores, the across-school differences are larger, but 53% of the variance remains within-school. One possible explanation for this difference may be that higher turnover rates at certain schools are driving the increase in across-school variation for this teacher quality measure. Similar patterns are found for ELA, though the distribution for VAM is more evenly distributed than math VAMs (53.7% and 74.4% due to within-school variation, respectively).

Tables 2 and 3 show that overall, there are TQGs by school-level disadvantage and, for SWD in general education classrooms, significant differences by disability status in average teacher VAMs, teacher evaluation scores, and hiring scores. However, these findings are unable to shed light on how these two factors interact. The rest of the paper explores how teacher quality varies when we examine student disability status and school poverty levels simultaneously.

Research Question 3: Do SWD vs non-SWD gaps vary by school-level disadvantage?

Table 4 presents the mean teacher quality scores and quality gaps across disability status and school disadvantage, with each column representing a different teacher characteristic of interest. Columns (1)-(4) present the results for math teachers while columns (5)-(8) present our findings for ELA teachers. Panel A presents the results for the most advantaged (<70% FRL) bin. We find significant SWD versus non-SWD TQGs for math VAM (-0.047 SD) and ELA teacher evaluation score (-0.085 SD).

Panels B and C present our findings for the middle (70-<95% FRL) and most disadvantaged (>=95% FRL) schools, respectively. As in panel A, we see that SWDs have teachers with lower math VAMs (Panel B: -0.014, Panel C: -0.019), though this difference is only significant for the highest poverty schools. We find no statistically significant math TQGs based on the other teacher characteristics. In ELA, we find SWDs in the middle FRL bin are exposed to teachers with lower evaluation scores (-0.072 SD) and are more likely to have teachers with two or less years of experience (0.8 percentage point difference) than their non-SWD peers. Aside from an increased likelihood of having a novice teacher (0.7 percentage points), SWDs in the highest FRL bin are exposed to ELA teachers with similar qualities as their non-SWD peers. In subgroup analyses (Appendix Tables 2A-D), we find that much of the TQGs are being driven by middle school grades (sixth through eighth grades) rather than earlier grades.

The last row of Panels B and C displays the p-values for the disadvantage gap when we compare TQGs from each bin to the lowest FRL bin. Our estimates suggest that, overall, TQGs are similar across FRL bins. The one exception is a significantly larger ELA teacher evaluation TQG in the lowest FRL bin (-0.085) compared to the highest FRL bin (-0.019).

Appendix Table 3 presents the between- and within-school variance decomposition within disadvantage bins. In general, the ratio of within- to between-school variance is similar across bins, though there are differing patterns for novice teachers. Within the lowest poverty schools, TQGs for novice math teachers are mostly driven by within-school sorting (96%), while TQGs in the highest poverty schools are more evenly split. For ELA, novice teacher disparities are mostly driven by within school differences for the middle FRL bin (73.6%), while differences in both the lowest and highest FRL bins are more equally split by across and within school sorting.

Research Question 4: Do TQGs across school-level disadvantage vary by specific disability type?

Looking across all students with disabilities may mask heterogeneity by disability type. Consequently, we disaggregate our data to examine the three largest disability subgroups: specific learning disability (SLD, ~57% of SWDs), autism (~11% of SWDs), and speech/language impairment (SLI, ~23% of SWDs). Table 5 presents our math TQG subsample estimates. Within each panel, we present the TQG (for each specific disability compared to non-SWD), standard errors, and sample size for each cell. Following the format in Table 4, we also include p-values for disadvantage gaps, which measure if TQGs in each FRL bin are significantly different from the TQG in the most advantaged schools (<70% FRL).

Results for students with SLD follow a similar pattern to the overall sample. Across the lowest and highest FRL bins, SLD students have teachers with significantly lower math VAMs (-0.074 and -0.029 SD, respectively) compared to their peers without disabilities. Our estimates also suggest that SLD students tend to have teachers with lower evaluation scores (significant for the middle and highest FRL bins, -0.087 SD and -0.053 SD, respectively). Like our main findings, we find no significant differences in terms of hiring scores and novice teachers. Interestingly, we find that the VAM TQGs in the lowest poverty bin is significantly greater than the VAM TQGs in the more disadvantaged school groups. TQGs across other teacher quality measures did not vary across FRL bins, suggesting little correlation between school-level disadvantage and SWD TQGs.

Estimates for students with autism and SLI suggest few significant differences from their non-SWD peers. If anything, our estimates suggest that depending on the FRL bin, these subgroups may be accessing higher quality teachers than their non-SWD peers. For example, students with SLI in the middle and highest FRL bins have teachers with significantly higher VAMs (0.076 and 0.058 SD, respectively) than their non-SWD peers. Additionally, SLI in the highest FRL schools are more likely to access teachers with significantly higher evaluation scores (0.076 SD) and less likely to have novice teachers. Overall, we generally do not find evidence TQGs that vary by school-level disadvantage. The one exception is that the SLI teacher evaluation TQG between the most disadvantaged schools (0.076 SD) is significantly larger than the SLI TQG in the lowest FRL schools (-0.021 SD).

Appendix Table 4 displays the estimates for ELA teachers. Like the findings in Table 5, SLD students followed the same pattern for overall ELA teacher quality differences. We find no evidence of teacher quality gaps for students with autism. Similar to the math results, we find evidence that SLI students are more likely to be exposed to higher quality teachers than their non-SWD peers. Of note, SLI students are consistently exposed to higher ELA VAM teachers across all FRL bins (ranging from 0.207 to 0.308 SD).

Discussion & Policy Implications

In this study, we provide some of the first evidence documenting the extent of teacher quality gaps between students with and without disabilities, as well as differences in TQGs by school-level disadvantage. While previous research has found few TQGs between SWDs and their peers, we find significant TQGs for some measures of teacher quality. Specifically, we see significant TQGs for math VAMs and ELA teacher evaluation scores, and small but statistically significant gaps for exposure to novice ELA teachers. We also extend the current literature by showing that these TQGs do not generally increase with school-level disadvantage. Additionally, we explore TQGs by disability subgroup, finding evidence that TQGs are concentrated within students with specific learning disabilities, the largest subgroup of SWDs. Students with autism or speech/language impairment do not seem to be placed with teachers of lower quality than their peers. Our findings for students with speech impairments are consistent with Gilmour and Henry (2018), though our TQGs for students with LD stand in contract to theirs. One potential explanation for this may lie in local differences in how mild disabilities are categorized (e.g. Saatcioglu & Skrtic, 2019), highlighting the importance of studying special education in different contexts such as urban areas like Los Angeles.

Existing literature from general education shows that schools tend to assign novice or less-effective teachers to classes with larger proportions of low-performing students (e.g. Bruno et al, 2019; Kalogrides et al, 2013; Lankford et al, 2002). In the case of SWDs, it does not appear that these students are being actively sorted into math classrooms based on observable teacher characteristics, as we only find significant gaps for VAMs. Instead, we believe that principals may be sorting on unobservable (to us) characteristics that are highly correlated with math VAMs. For ELA teachers, our findings suggest that while there may be some sorting across observable teacher characteristics, these do not result in exposure to lower quality ELA teachers as measured by VAMs.

Our variance decomposition suggests that the majority of SWD TQGs in LAUSD are due to within- rather than between-school factors, and is in line with previous research on SWD TQGs (Gilmour & Henry, 2018). For practitioners, the implication is that solutions to the SWD quality gaps do not necessarily have to come from district and state policies aimed at recruiting and retaining higher quality teachers overall—though these avenues can certainly help schools obtain more high-quality teachers. Instead, our estimates suggest that a more immediate solution could be to shift student compositions amongst existing teachers within schools, particularly if schools are trying to adhere to the *Endrew F.* court decision to ensure equitable outcomes for SWDs.

Our evidence contributes to a growing literature addressing the quality of instruction received by SWDs and suggests several avenues for future research. The field would benefit from qualitative interviews to explore what, if any, factors principals take into consideration when matching students to teachers, and how this varies by school level and subject area. It is probable that teacher characteristics beyond those in the current study are used to determine how students are matched to teachers. For example, principals may pair certain SWDs with a teacher who is particularly strong at engaging their students in classroom activities or who have strong classroom management skills. Furthermore, these traits may play an important role for improving SWDs' academic outcomes. Similarly, future work could unpack the instructional practices of general education teachers who are comparatively more effective at improving outcomes for SWDs. This empirical work could help practitioners move towards the end goal of more equitable academic outcomes for SWDs.

Endnotes

1. For example, we do not include students who attend home or hospital schools, special education centers, nor community day schools.
2. While it may also be of interest to examine TQGs for SWDs taught in special education classrooms by special education teachers, data limitations make this problematic. In particular, we can only calculate VAM scores for approximately 15% of special education teachers (SETs) in our sample because students' previous test scores are used to construct VAMs and few SWDs with SETs have valid test scores from the previous year. Additionally, while special education teachers are evaluated on the same observation instrument and hiring criteria as general education teachers, researchers and practitioners argue that these shared measures should not be used to measure special education teacher quality since special education teachers' work responsibilities and preparation programs are different from those for general education teachers (see Brownell, Ross, Colon, & McCallum, 2005 for a review). Furthermore, recent research on teacher evaluations suggest that SETs may systematically receive lower evaluation scores since effective teaching looks different for special education teachers than general education teachers--particularly given the individualized nature of special education (Johnson and Semmelroth, 2013; Jones and Brownell, 2014). For completeness, we calculate non-VAM TQGs for students in special education classrooms (compared to their peers in general education classrooms), broken down by school disadvantage level and disability type, in Appendix Tables 8 and 9. Given these concerns we are hesitant to

say whether these results are indicative of the existence or non-existence of quality gaps among SWDs with special education teachers compared to their non-SWD peers.

3. While typically IDEA only requires districts to designate a primary disability, along with blindness and deafness as secondary disabilities, LAUSD operates under a consent decree that requires more detailed tracking (Weintraub, Myers, Hehir, Jaque-Anton, 2008). For our main analysis, we are focused on whether students have any disabilities listed. Consequently, we do not separately account for students with multiple disabilities. In disability-specific analysis, we include any students who have that disability subcategory listed in their individualized education plan (including those with multiple disabilities). We have also run analysis that excludes students with multiple disabilities and find little difference. Results are available upon request.
4. In analysis not shown, we also disaggregate schools by share of students who are underrepresented minorities or have low prior test scores. The results are qualitatively similar and available upon request.
5. As a sensitivity check, we have also split schools into four bins (<70, 70-<95, 95-<0.978, >=0.978), and five bins (<70, 70-<80, 80-<90, 90-<95, >=95). Results are similar to those found in our main tables and available upon request.
6. Only grades 3-8 have test scores that are usable in standard Value Added Measures. Consequently, we calculate VAMs only for students in grades 4-8 (leaving out grade 3 to ensure there is a lagged score). All VAM scores are standardized across the full sample of LAUSD teachers in each year (instead of just our study sample). Accordingly, the overall VAM mean (displayed in Table 1) is not zero. The teacher quality literature has used multiple different ways to measure teacher value-added. As a robustness check, we also

estimate one-year teacher value-added measures that use teacher fixed effects and includes student- and classroom-level demographics (see Appendix B for more details about the construction of these models). To address concerns that student characteristics are endogenous to teacher value-added in time *t*, we use teacher's value-added score in t-1 as a measure for teacher quality in time *t*. Additionally, we create an alternative VAM score for teachers that exclude students with disabilities from VAM calculations. These results are presented in Appendix Table 5 and similar to the ones we show in our main tables.

7. We include four types of testing accommodation flags: technology (i.e. text-to-speech software), setting (i.e. small group setting), time (i.e. extended time), and format (i.e. streamlined version of text).
8. Teacher evaluation may be deferred for employees with ten or more years of satisfactory service, have not received a "notice of unsatisfactory act of service" in the past four years, and had fewer than 13 unprotected absences in the past year.
9. We also analyzed results by teachers' final evaluation score, which only has three values: below standard performance, meets standard performance, and exceeds standard performance. Since less than 5% of teachers each year do not pass the evaluation, there is not much variation in the final score. Consequently, we focus our main results on the average score across all subcomponents. The average score across all subcomponents does not necessarily map onto the final evaluation score (though it very rarely does not match) and has the additional benefit of having more variation to distinguish between teacher scores. Results for final evaluation scores available upon request. We have also analyzed a few alternative measures for teacher evaluation scores. Following Kraft et al

(2018), we create a measure of overall performance using a graded response model for all subcomponents, *theta*, as well as a residualized *theta* measure that removes classroom- and school-level student demographic variation. However, since teachers are not all assessed on the same components, we also create a *theta* based only on the three subcomponents that are mandatory for all teachers, as well as a residualized *theta* score based on these three subcomponents. Finally, we also individually analyze the raw scores for each mandatory subcomponent. Across all these differing teacher evaluation measures, we find little evidence of SWD vs non-SWD teacher quality gaps. All results are displayed in Appendix Tables 6 and 7.

10. As a sensitivity check, we also define "novice teacher" as those with five or fewer years of experience. Results are qualitatively similar.
11. We also present results on student- and school-level demographics in Appendix Table 1. Similar to previous literature, we find that as school-level disadvantage increases, so does the share of students who are black, Hispanic, or labeled as English Language Learner.
12. We do not present results for the "other disabilities" subgroup since this group encompasses a large range of disabilities from emotional disturbance to intellectual disability and interpreting any potential gaps would be difficult. However, for completeness, we include this indicator variable in our VAM calculations.

Citations

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, *25*(1), 95-135.

Angrist, J. D., & Guryan, J. (2008). Does teacher testing raise teacher quality? Evidence from state certification requirements. *Economics of Education Review*, *27*(5), 483–503.

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.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2008). Consequences? The Impact of Assessment and Accountability on Teacher Recruitment and Retention Are There Unintended Consequences? *Public Finance Review*, *36*, 88–111. https://doi.org/10.1177/1091142106293446

Brownell, M. T., Ross, D. D., Colón, E. P., & McCallum, C. L. (2005). Critical features of special education teacher preparation: A comparison with general teacher education. *The Journal of Special Education*, *38*(4), 242-252.

Bruno, P. and Strunk, K. O. (2019). Making the Cut: The Effectiveness of Teacher Screening and Hiring in the Los Angeles Unified School District. *Educational Evaluation and Policy Analysis*, *41*(4), 426-260.

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.

Child, A. E., Cirino, P. T., Fletcher, J. M., Willcutt, E. G., & Fuchs, L. S. (2019). A cognitive dimensional approach to understanding shared and unique contributions to reading, math, and attention skills. *Journal of Learning Disabilities*, *52*(1), 15-30.

Chingos, M. M., & Peterson, P. E. (2011). It's easier to pick a good teacher than to train one: Familiar and new results on the correlates of teacher effectiveness. *Economics of Education Review*, *30*(3), 449–465.

Chudowsky, N., Chudowsky, V., and Keber, N. (2009). *State Test Score Trends Through 2007–08, Part 4: Has Progress Been Made in Raising Achievement for Students with Disabilities?* Washington, DC: Center on Education Policy.

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

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review*, *26*(6), 673–682.

Endrew F. v. Douglas County School District RE–1, 580 U.S. ___ (2017)

Fuchs, L. S., Geary, D. C., Fuchs, D., Compton, D. L., & Hamlett, C. L. (2016). Pathways to third-grade calculation versus word-reading competence: Are they more alike or different?. *Child Development*, *87*(2), 558-567.

Gilmour, A., & Henry, G. (2018). A comparison of teacher quality in math for late elementary and middle school students with and without disabilities. *The Elementary School Journal*, *118*(3), 426–451.

Gilmour, A. F., & Henry, G. T. (2020). Who Are the Classmates of Students With Disabilities in Elementary Mathematics Classrooms?. *Remedial and Special Education*, *41*(1), 18-27.

Gilmour, A. F., & Wehby, J. H. (2019). The association between teaching students with disabilities and teacher turnover. *Journal of Educational Psychology*. doi:10.1037/edu0000

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., Quince, V., & Theobald, R. (2018). Has it always been this way? Tracing the evolution of teacher quality gaps in US public schools. *American Educational Research Journal*, *55*(1), 171-201.

Hanushek, E. A., Kane, T.J., & Rivkin, S.G. (2004). Why Public Schools Lose Teachers. *The Journal of Human Resources*, *39*(2), 362-354.

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. https://doi.org/10.1080/2330443X.2016.1182878

Johnson, E., & Semmelroth, C. L. (2014). Special Education Teacher Evaluation. *Assessment for Effective Intervention*, *39*(2), 71–82.

Jones, N. D. & Brownell, M. (2013a). Examining the Use of Classroom Observations in the Evaluation of Special Education Teachers. *Assessment for Effective Intervention, 39*, 112-124.

Jones, N. D., Buzick, H., and Turkan, S. (2013b). Including Students with Disabilities and English Learners in Measures of Educator Effectiveness. *Educational Researcher*, *42*, 234-241.

Kalogrides, D., & Loeb, S. (2013). Different Teachers, Different Peers: The Magnitude of Student Sorting Within Schools. *Educational Researcher*, *42*(6), 304–316.

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review*, *27*(6), 615-631.

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

Ladd, H. F., & Sorensen, L. C. (2017). Returns to teacher experience: Student achievement and motivation in middle school. *Education Finance and Policy*, *12*(2), 241-279.

Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Education Evaluation and Policy Analysis*, *24*(1), 37-62.

National Center for Education Statistics. (2015). *The Condition of Education*. Retrieved from https://nces.ed.gov/pubs2015/2015144.pdf.

Papay, J. P., & Kraft, M. A. (2015). Productivity returns to experience in the teacher labor market: Methodological challenges and new evidence on long-term career improvement. *Journal of Public Economics*, *130*, 105-119.

Rice, J. K. (2010). The Impact of Teacher Experience: Examining the Evidence and Policy Implications. Brief No. 11. National center for analysis of longitudinal data in education research.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417-458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, *94*(2), 247-252.

Saatcioglu, A., & Skrtic, T. M. (2019). Categorization by organizations: Manipulation of disability categories in a racially desegregated school district. *American Journal of Sociology*, *125*(1), 184-260.

Sass, T. R., Hannaway, J., Xu, Z., Figlio, D. N., & Feng, L. (2012). Value added of teachers in high-poverty schools and lower poverty schools. *Journal of Urban Economics*, *72*(2-3), 104-122.

Schulte, A. C., & Stevens, J. J. (2015). Once, Sometimes, or Always in Special Education Mathematics Growth and Achievement Gaps. *Exceptional Children*, *81*, 370-387.

Schulte, A. C., Elliott, J. J., Tindal, S. N., & Nese, G. (2016). Achievement Gaps for Students with Disabilities: Stable, Widening, or Narrowing on a State-wide Reading Comprehension Test? *Journal of Educational Psychology*, *16*, 925–942.

Staiger, D. O., & Rockoff, J. E. (2010). Searching for effective teachers with imperfect information. *Journal of Economic Perspectives*, *24*(3), 97-118.

Swaak, Taylor. (2020, January 22). For the first time in more than 20 years, LAUSD is in full control of its special ed system. As parents worry about accountability, the district shifts its focus. *The New York Times*, Retrieved from http://laschoolreport.com/for-the-first-time-in-20-years-lausd-is-in-full-control-of-its-special-ed-system-as-parents-worry-about-accountability-the-district-shifts-its-focus/

U.S. Department of Education. (2019). *The NCES Fast Facts Tool provides quick answers to many education questions.* Washington, DC: National Center for Education Statistics.

Weintraub, F. J., Myers, R. M., Hehir, T., & Jaque-Anton, D. (2008). A Contextual Overview of the Modified Consent Decree in the Los Angeles Unified School District. *Journal of Special Education Leadership*, *21*(2), 51-57.

Table 1. Teacher Characteristics in Math Classes by Teacher Sample and FRL School Bins

	Panel A. VAM Sample				Panel B. Teacher Eval Sample				Panel C. Hiring Score Sample				Panel D. Novice Sample			
FRL Group	**<=70%**	**70-95%**	**>=95%**	**Overall**	**<=70%**	**70-95%**	**>=95%**	**Overall**	**<=70%**	**70-95%**	**>=95%**	**Overall**	**<=70%**	**70-95%**	**>=95%**	**Overall**
No. of Schools	121	133	341	595	121	128	351	600	106	98	284	488	123	134	362	619
General Education Teacher Characteristics																
No. of Teachers	836	993	3,121	4,950	983	1,214	4,094	6,291	257	217	785	1,259	2,423	2,477	8,207	13,107
No. of Teacher-Years	2,764	3,398	10,253	16,415	1,783	2,338	7,749	11,870	593	512	1,783	2,888	7,519	8,087	26,226	41,832
Mean Experience	9.41	9.27	9.31	9.32	8.82	8.92	9.11	9.02	4.65	4.81	5.14	4.98	9.36	9.36	9.37	9.36
%Novice	0.03	0.03	0.03	0.03	0.03	0.03	0.02	0.02	0.34	0.31	0.29	0.31	0.03	0.03	0.03	0.03
%Masters or Higher	0.34	0.40	0.39	0.38	0.34	0.37	0.36	0.36	0.39	0.39	0.36	0.37	0.34	0.36	0.36	0.36
%Fem	0.72	0.68	0.66	0.68	0.78	0.76	0.73	0.74	0.90	0.80	0.85	0.85	0.80	0.75	0.75	0.76
%White	0.53	0.36	0.25	0.33	0.52	0.31	0.22	0.29	0.62	0.28	0.25	0.33	0.53	0.35	0.23	0.31
%Black	0.06	0.10	0.10	0.09	0.06	0.09	0.09	0.08	0.02	0.05	0.09	0.07	0.06	0.10	0.09	0.08
%Hispanic	0.19	0.32	0.53	0.42	0.22	0.40	0.58	0.48	0.19	0.41	0.51	0.42	0.20	0.36	0.56	0.45
%Have VAM	1.00	1.00	1.00	1.00	0.61	0.61	0.56	0.58	0.46	0.63	0.57	0.56	0.53	0.58	0.51	0.52
Mean VAM	-0.02	-0.14	-0.05	-0.07	-0.11	-0.19	-0.13	-0.14	-0.04	-0.20	-0.08	-0.10	-0.02	-0.14	-0.05	-0.07
%Have Teacher Eval	0.29	0.30	0.33	0.32	1.00	1.00	1.00	1.00	0.57	0.61	0.57	0.58	0.25	0.29	0.30	0.29
Mean Teacher Eval	0.35	0.14	0.11	0.16	0.36	0.19	0.16	0.20	0.28	0.13	0.16	0.18	0.36	0.19	0.16	0.20
%Have Hiring Score	0.06	0.07	0.08	0.08	0.17	0.15	0.13	0.14	1.00	1.00	1.00	1.00	0.07	0.07	0.07	0.07
Mean Hiring Score	0.10	0.08	-0.08	-0.01	0.18	0.06	-0.03	0.03	0.13	0.06	-0.04	0.01	0.13	0.06	-0.04	0.01

Observations are at student-teacher cell level pooled across school years 2014-15 to 2017-18. Each column represents teacher characteristics in schools binned by the percent of FRL eligible students. A school's FRL bin is defined by taking a three year average of the percent of FRL eligible students. Experience represents years of experience and is top-coded at 10 years. Novice is defined by any teacher with fewer than 2 years of experience. VAM, Teacher Eval, and Hiring Score are z-scored measures for value-added, evaluation scores, and hiring scores, respectively. %Have indicates what percent of the given sample has a VAM, Teacher Eval, or Hiring Score measure. Panels A, B, C, and D represent the VAM, Teacher Eval, Hiring Score, and Novice samples, respectively. Panel D also represents the population sample.

Table 2. Average Teacher Quality, by Subject and FRL Bin								
	Math				ELA			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	VAM	Teach Eval	Hiring Score	Novice	VAM	Teach Eval	Hiring Score	Novice
A. <70% FRL Teacher Quality	-0.016	0.361	0.123	0.034	-0.12	0.323	0.232	0.032
Std. Dev	(0.529)	(0.729)	(1.084)	(0.180)	(0.908)	(0.773)	(0.875)	(0.176)
n	118,370	55,682	16,546	223,871	131,167	56,593	17,853	235,623
B. 70% - <95% FRL Teacher Quality	-0.135	0.19	0.061	0.03	-0.367	0.202	0.159	0.031
Std. Dev	(0.620)	(0.826)	(0.710)	(0.172)	(1.095)	(0.852)	(0.888)	(0.175)
n	141,686	69,980	16,777	245,351	148,669	72,322	19,787	250,140
Disadv. Gap	-0.119*	-0.171**	-0.062	-0.004	-0.247*	-0.121*	-0.073	-0.001
Disadv. Gap [p-value]	[0.023]	[0.01]	[0.48]	[0.85]	[0.048]	[0.045]	[0.499]	[0.822]
C. >= 95% FRL Teacher Quality	-0.055	0.161	-0.039	0.031	-0.23	0.172	-0.01	0.036
Std. Dev	(0.617)	(0.854)	(0.928)	(0.173)	(1.126)	(0.865)	(0.916)	(0.186)
n	357,985	213,663	49,676	708,089	371,134	216,808	60,273	715,865
Disadv. Gap	0.039	0.200***	0.162	0.003	0.11	0.151***	0.242**	-0.004
Disadv. Gap [p-value]	[0.289]	[0.00]	[0.114]	[0.754]	[0.257]	[0.001]	[0.004]	[0.14]

Disadv. Gap represents difference in Teacher Quality of FRL bin from the least disadvantaged bin (<70% FRL). Observations are at student-teacher cell level pooled across school years 2014-15 to 2017-18. A school's FRL bin is defined by taking a four year average of the percent of FRL eligible students. Novice is defined by any teacher with fewer than 2 years of experience. VAM (value-added measure), Teacher Eval, and Hiring Score are z-scored measures. *p<0.05, **p<0.01, ***p<0.001

Table 3. SWD vs non-SWD Teacher Quality Gaps				
	(1)	(2)	(3)	(4)
	VAM	Teach Eval	Hiring Score	Novice
Panel A. Math				
SWD Teacher Quality	-0.088	0.174	-0.002	0.033
Non-SWD Teacher Quality	-0.064	0.202	0.015	0.031
Teacher Quality Gap	-0.024***	-0.028*	-0.017	0.002
Std. Error	(0.007)	(0.011)	(0.021)	(0.001)
SWD *n*	50,521	27,276	6,706	91,062
non-SWD *n*	567,520	312,049	76,293	1,086,249
between variance	*0.253*	*0.341*	*0.475*	*0.332*
within variance	*0.747*	*0.659*	*0.525*	*0.668*
Panel B. ELA				
SWD Teacher Quality	-0.260	0.164	0.084	0.040
Non-SWD Teacher Quality	-0.237	0.207	0.067	0.034
Teacher Quality Gap	-0.023	-0.042***	0.017	0.006***
Std. Error	(0.015)	(0.013)	(0.024)	(0.001)
SWD *n*	53,604	28,401	8,940	93,437
non-SWD *n*	597,366	317,322	88,973	1,108,191
between variance	*0.463*	*0.34*	*0.45*	*0.406*
within variance	*0.537*	*0.66*	*0.55*	*0.594*

Between/Within variance calculated from Eq (1) with the addition of school fixed-effects. These represent the variance of teacher quality gaps between and within schools. Standard errors clustered at school level. SWD stands for Students with Disabilities. Teacher Quality Gap represents the difference between the SWD and non-SWD exposure rates. Observations are at student-teacher cell level pooled across school years 2014-15 to 2017-18. Novice is defined by any teacher with fewer than 2 years of experience. VAM (value-added measure), Teacher Eval, and Hiring Score are z-scored measures. *p<0.05, **p<0.01, ***p<0.001

Table 4. SWD with GET vs non-SWD Teacher Quality Gaps by Subject and FRL Bins

	Math				ELA			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	VAM	Teach Eval	Hiring Score	Novice	VAM	Teach Eval	Hiring Score	Novice
A. <70% FRL								
TQG	-0.047***	-0.028	-0.019	0.000	0.001	-0.085**	0.032	0.001
Std. Error	(0.014)	(0.026)	(0.050)	(0.003)	(0.024)	(0.028)	(0.042)	(0.003)
SWD *n*	9,258	4,397	1,266	16,865	9,839	4,318	1,467	17,366
non-SWD *n*	109,112	51,285	15,280	207,006	121,328	52,275	16,386	218,257
B. 70% - <95% FRL								
TQG	-0.014	-0.041	-0.001	0.007	-0.012	-0.072*	0.084	0.008*
Std. Error	(0.018)	(0.024)	(0.035)	(0.004)	(0.037)	(0.030)	(0.068)	(0.004)
SWD n	12,152	5,913	1,429	20,005	13,112	6,294	1,915	20,751
non-SWD n	129,534	64,067	15,348	225,346	135,557	66,028	17,872	229,389
Disadv. Gap [p-value]	[0.143]	[0.71]	[0.757]	[0.19]	[0.76]	[0.752]	[0.521]	[0.163]
C. >= 95% FRL								
TQG	-0.019*	-0.023	-0.021	0.001	-0.027	-0.019	-0.002	0.007***
Std. Error	(0.008)	(0.014)	(0.029)	(0.001)	(0.020)	(0.015)	(0.028)	(0.002)
SWD n	29,111	16,966	4,011	54,192	30,653	17,789	5,558	55,320
non-SWD n	328,874	196,697	45,665	653,897	340,481	199,019	54,715	660,545
Disadv. Gap [p-value]	[0.083]	[0.879]	[0.974]	[0.865]	[0.368]	[0.037]	[0.487]	[0.111]

TQG stands for Teacher Quality Gap. GET stands for General Education Teacher. SWD stands for Students with Disabilities. Disadv. Gap represents how similar the Free-Reduced Lunch bins are from the least disadvantaged bin (<70% FRL). Observations are at student-teacher cell level pooled across school years 2014-15 to 2017-18. A school's FRL bin is defined by taking a four year average of the percent of FRL eligible students. Exp represents years of experience and is top-coded at 10 years. Novice is defined by any teacher with fewer than 2 years of experience. VAM (value-added measure), Teacher Eval, and Hiring Score are z-scored measures.

Table 5. Math Teacher Quality Gaps by Disability Type (vs. No Disability) and FRL Bins				
SWD with GET vs non-SWD				
	(1)	(2)	(3)	(4)
Disability & FRL Group	VAM	Teach Eval	Hiring Score	Novice
A. Specific Learning				
<70% FRL TQG	-0.074***	-0.049	-0.062	0.002
Std. Error	(0.019)	(0.047)	(0.075)	(0.004)
SWD n	5,071	2,110	560	7,903
non-SWD n	109,112	51,285	15,280	207,006
70-<95% FRL TQG	-0.022	-0.087*	-0.02	0.01
Std. Error	(0.019)	(0.038)	(0.044)	(0.006)
SWD n	8,078	3,429	859	11,935
non-SWD n	129,534	64,067	15,348	225,346
Disadv. Gap [p-value]	[0.051]	[0.519]	[0.627]	[0.252]
>=95% FRL TQG	-0.029**	-0.053**	-0.016	0.004
Std. Error	(0.009)	(0.020)	(0.040)	(0.002)
SWD n	19,997	9,930	2,446	31,959
non-SWD n	328,874	196,697	45,665	653,897
Disadv. Gap [p-value]	[0.032]	[0.928]	[0.592]	[0.661]
B. Autism				
<70% FRL TQG	0	0.012	-0.006	-0.004
Std. Error	(0.021)	(0.035)	(0.081)	(0.004)
SWD n	1,405	715	199	2,834
non-SWD n	109,112	51,285	15,280	207,006
70-<95% FRL TQG	0.01	0.063	0.053	0.004
Std. Error	(0.032)	(0.041)	(0.047)	(0.005)
SWD n	1,126	641	139	2,090
non-SWD n	129,534	64,067	15,348	225,346
Disadv. Gap [p-value]	[0.784]	[0.338]	[0.529]	[0.226]
>=95% FRL TQG	0.005	0.033	-0.034	-0.003
Std. Error	(0.016)	(0.026)	(0.056)	(0.003)
SWD n	2,267	1,405	325	4,632
non-SWD n	328,874	196,697	45,665	653,897
Disadv. Gap [p-value]	[0.851]	[0.619]	[0.779]	[0.853]
C. Speech/Language				
<70% FRL TQG	0.022	-0.021	0.153	-0.004
Std. Error	(0.031)	(0.039)	(0.089)	(0.004)
SWD n	888	821	260	3,098
non-SWD n	109,112	51,285	15,280	207,006
70-<95% FRL TQG	0.076*	0.056	-0.03	-0.002
Std. Error	(0.029)	(0.039)	(0.070)	(0.004)
SWD n	1,223	1,137	231	3,647
non-SWD n	129,534	64,067	15,348	225,346
Disadv. Gap [p-value]	[0.207]	[0.158]	[0.107]	[0.752]
>=95% FRL TQG	0.058***	0.076***	0.002	-0.006**
Std. Error	(0.014)	(0.023)	(0.042)	(0.002)
SWD n	3,821	4,249	855	13,701
non-SWD n	328,874	196,697	45,665	653,897
Disadv. Gap [p-value]	[0.293]	[0.03]	[0.125]	[0.706]

TQG stands for Teacher Quality Gap. GET stands for General Education Teacher. SWD stands for Students with Disabilities. Standard errors clustered at school level. Disadv. Gap represents how similar the Free-Reduced Lunch bins are from the least disadvantaged bin (<70% FRL TQG FRL). Observations are at student-teacher cell level pooled across school years 2014-15 to 2017-18. A school's FRL bin is defined by taking a four year average of the percent of FRL eligible students. Novice is defined by any teacher with fewer than 2 years of experience. VAM (value-added measure), Teacher Eval, and Hiring Score are z-scored measures. *p<0.05, **p<0.01, ***p<0.001

Appendix Table 1. Student Characteristics in Math Classes by Teacher Sample and FRL School Bins

	Panel A. VAM Sample				Panel B. Teacher Eval Sample				Panel C. Hiring Score Sample				Panel D. Novice Sample			
FRL Group	**<=70%**	**70-95%**	**>=95%**	**Overall**	**<=70%**	**70-95%**	**>=95%**	**Overall**	**<=70%**	**70-95%**	**>=95%**	**Overall**	**<=70%**	**70-95%**	**>=95%**	**Overall**
Student-Years (*n*)	118,278	141,555	357,625	617,458	55,621	69,898	213,237	338,756	16,528	16,706	49,481	82,715	223,650	245,041	706,975	1,175,666
Student with Disability	0.078	0.086	0.081	0.082	0.079	0.084	0.079	0.080	0.077	0.085	0.081	0.081	0.075	0.082	0.077	0.077
Female	0.502	0.497	0.501	0.500	0.499	0.489	0.499	0.497	0.501	0.487	0.499	0.497	0.500	0.495	0.499	0.498
FRL Eligible	0.466	0.837	0.975	0.846	0.479	0.865	0.983	0.876	0.397	0.869	0.981	0.842	0.417	0.845	0.976	0.843
English Language Learner	0.152	0.240	0.326	0.273	0.154	0.298	0.391	0.333	0.157	0.309	0.407	0.337	0.159	0.298	0.411	0.340
White	0.343	0.068	0.020	0.093	0.338	0.072	0.025	0.086	0.395	0.069	0.029	0.110	0.386	0.072	0.025	0.104
Black	0.096	0.101	0.052	0.072	0.101	0.085	0.061	0.072	0.101	0.101	0.090	0.095	0.094	0.095	0.058	0.072
Hispanic	0.396	0.738	0.903	0.768	0.404	0.754	0.891	0.783	0.335	0.727	0.856	0.726	0.354	0.741	0.893	0.759
Other Race	0.165	0.092	0.025	0.067	0.157	0.090	0.023	0.059	0.169	0.103	0.025	0.070	0.167	0.092	0.025	0.066
Previous Math Score	0.618	0.025	-0.170	0.031	0.575	0.031	-0.179	-0.002	0.582	0.016	-0.205	-0.019	0.613	0.022	-0.172	0.025
%Have Prev. Math Score	0.920	0.918	0.879	0.896	0.581	0.580	0.530	0.548	0.523	0.680	0.608	0.606	0.532	0.585	0.501	0.525
Schools (*n*)	121	133	341	595	121	128	351	600	106	98	284	488	123	134	362	619
Student with Disability	0.102	0.127	0.124	0.120	0.102	0.126	0.121	0.118	0.101	0.126	0.123	0.119	0.102	0.129	0.121	0.119
Female	0.497	0.497	0.490	0.493	0.497	0.491	0.490	0.492	0.498	0.481	0.486	0.487	0.500	0.496	0.488	0.492
FRL Eligible	0.414	0.850	0.976	0.833	0.414	0.847	0.976	0.835	0.400	0.847	0.976	0.825	0.420	0.851	0.976	0.838
English Language Learner	0.158	0.306	0.417	0.340	0.158	0.314	0.430	0.350	0.160	0.311	0.436	0.351	0.156	0.306	0.428	0.347
White	0.397	0.082	0.026	0.114	0.397	0.085	0.026	0.113	0.416	0.077	0.027	0.121	0.392	0.082	0.026	0.111
Black	0.107	0.128	0.072	0.092	0.107	0.125	0.074	0.092	0.101	0.130	0.077	0.093	0.111	0.131	0.073	0.093
Hispanic	0.340	0.709	0.876	0.730	0.340	0.708	0.873	0.731	0.328	0.712	0.869	0.720	0.344	0.708	0.876	0.734
Other Race	0.155	0.080	0.026	0.065	0.155	0.082	0.026	0.064	0.155	0.081	0.028	0.066	0.153	0.080	0.026	0.063
Previous Math Score	0.586	0.016	-0.171	0.026	0.586	0.022	-0.173	0.028	0.604	0.039	-0.179	0.038	0.586	0.009	-0.175	0.023

Each column represents teacher characteristics in schools binned by the percent of FRL eligible students. A school's FRL bin is defined by taking a three year average of the percent of FRL eligible students. Novice is defined by any teacher with fewer than 2 years of experience. VAM, Teacher Eval, and Hiring Score are z-scored measures for value-added, evaluation scores, and hiring scores, respectively. Panels A, B, C, and D represent the VAM, Teacher Eval, Hiring Score, and Novice samples, respectively. Panel D also represents the population sample.

Appendix Table 2A. Math Teacher Quality Gaps by FRL Bins, Grade 4-5 Only

SWD with GET vs non-SWD				
	(1)	(2)	(3)	(4)
	VAM	Teach Eval	Hiring Score	Novice
A. <70% FRL				
SWD Teacher Quality	0.064	0.341	0.341	0.034
Non-SWD Teacher Quality	0.097	0.389	0.278	0.036
TQG	-0.033**	-0.048*	0.063	-0.002
Std. Error	(0.012)	(0.021)	(0.040)	(0.003)
SWD n	4,640	2,896	920	11,805
non-SWD n	51,100	33,748	11,613	144,566
between variance	*0.235*	*0.354*	*0.466*	*0.041*
within variance	*0.765*	*0.646*	*0.534*	*0.959*
B. 70% - <95% FRL				
SWD Teacher Quality	0.072	0.295	0.025	0.027
Non-SWD Teacher Quality	0.057	0.297	0.038	0.024
TQG	0.015	-0.002	-0.014	0.003
Std. Error	(0.015)	(0.021)	(0.047)	(0.002)
SWD n	5,119	3,868	757	12,172
non-SWD n	54,458	42,317	8,472	142,830
between variance	*0.282*	*0.401*	*0.549*	*0.061*
within variance	*0.718*	*0.599*	*0.451*	*0.939*
Disadv. Gap [p-value]	[0.012]	[0.125]	[0.212]	[0.216]
C. >= 95% FRL				
SWD Teacher Quality	0.084	0.249	-0.015	0.03
Non-SWD Teacher Quality	0.086	0.243	-0.003	0.028
TQG	-0.002	0.006	-0.012	0.003*
Std. Error	(0.007)	(0.012)	(0.026)	(0.001)
SWD n	15,703	11,978	2,839	39,042
non-SWD n	183,225	144,479	32,857	489,650
between variance	*0.209*	*0.284*	*0.387*	*0.065*
within variance	*0.791*	*0.716*	*0.613*	*0.935*
Disadv. Gap [p-value]	[0.023]	[0.027]	[0.11]	[0.162]

TQG stands for Teacher Quality Gap. GET stands for General Education Teacher. SWD stands for Students with Disabilities. Between/Within variance calculated from Eq. (2) with the inclusion of school fixed-effects. Disadv. Gap represents how similar the FRL bins from the least disadvantaged bin (<70% FRL). Observations are at student-teacher cell level pooled across school years 2014-15 to 2016-17. A school's FRL bin is defined by taking a three year average of the percent of FRL eligible students. Novice is defined by any teacher with fewer than 2 years of experience. VAM, EDST, and Hire are z-scored measures for value-added, evaluation scores, and hiring scores, respectively. *p<0.05, **p<0.01, ***p<0.001

Appendix Table 2B. ELA Teacher Quality Gaps by FRL Bins, Grade 4-5 Only

SWD with GET vs non-SWD				
	(1)	(2)	(3)	(4)
	VAM	Teach Eval	Hiring Score	Novice
A. <70% FRL				
SWD Teacher Quality	0.253	0.34	0.332	0.035
Non-SWD Teacher Quality	0.277	0.386	0.267	0.037
TQG	-0.023	-0.046*	0.065	-0.001
Std. Error	(0.023)	(0.021)	(0.040)	(0.003)
SWD n	4,714	2,923	942	11,872
non-SWD n	51,923	33,949	11,834	145,037
between variance	*0.367*	*0.356*	*0.462*	*0.231*
within variance	*0.633*	*0.644*	*0.538*	*0.769*
B. 70% - <95% FRL				
SWD Teacher Quality	0.295	0.293	0.029	0.027
Non-SWD Teacher Quality	0.259	0.295	0.045	0.024
TQG	0.036	-0.001	-0.016	0.003
Std. Error	(0.029)	(0.021)	(0.046)	(0.002)
SWD n	5,220	3,873	767	12,236
non-SWD n	55,393	42,413	8,600	143,517
between variance	*0.374*	*0.401*	*0.546*	*0.06*
within variance	*0.626*	*0.599*	*0.454*	*0.94*
Disadv. Gap [p-value]	[0.113]	[0.129]	[0.182]	[0.258]
C. >= 95% FRL				
SWD Teacher Quality	0.213	0.248	-0.021	0.031
Non-SWD Teacher Quality	0.211	0.242	-0.006	0.028
TQG	0.002	0.006	-0.015	0.003*
Std. Error	(0.014)	(0.012)	(0.026)	(0.001)
SWD n	16,006	12,076	2,905	39,291
non-SWD n	187,111	145,292	33,701	492,496
between variance	*0.308*	*0.302*	*0.399*	*0.297*
within variance	*0.692*	*0.698*	*0.601*	*0.703*
Disadv. Gap [p-value]	[0.342]	[0.031]	[0.088]	[0.206]

TQG stands for Teacher Quality Gap. GET stands for General Education Teacher. SWD stands for Students with Disabilities. Between/Within variance calculated from Eq. (2) with the inclusion of school fixed-effects. Disadv. Gap represents how similar the FRL bins from the least disadvantaged bin (<70% FRL). Observations are at student-teacher cell level pooled across school years 2014-15 to 2016-17. A school's FRL bin is defined by taking a three year average of the percent of FRL eligible students. Novice is defined by any teacher with fewer than 2 years of experience. VAM, EDST, and Hire are z-scored measures for value-added, evaluation scores, and hiring scores, respectively. *p<0.05, **p<0.01, ***p<0.001

Appendix Table 2C. Math Teacher Quality Gaps by FRL Bins, Grades 6-8 only				
SWD with GET vs non-SWD				
	(1)	(2)	(3)	(4)
	VAM	Teach Eval	Hiring Score	Novice
A. <70% FRL				
SWD Teacher Quality	-0.184	0.322	-0.521	0.033
Non-SWD Teacher Quality	-0.109	0.312	-0.359	0.027
TQG	-0.074**	0.011	-0.162	0.006
Std. Error	(0.024)	(0.063)	(0.125)	(0.006)
SWD n	4,618	1,501	346	5,060
non-SWD n	58,012	17,537	3,667	62,440
between variance	*0.432*	*0.542*	*0.383*	*0.474*
within variance	*0.568*	*0.458*	*0.617*	*0.526*
B. 70% - <95% FRL				
SWD Teacher Quality	-0.309	-0.118	0.103	0.053
Non-SWD Teacher Quality	-0.272	-0.008	0.089	0.04
TQG	-0.037	-0.111	0.013	0.013
Std. Error	(0.022)	(0.057)	(0.056)	(0.009)
SWD n	7,033	2,045	672	7,833
non-SWD n	75,076	21,750	6,876	82,516
between variance	*0.325*	*0.548*	*0.734*	*0.462*
within variance	*0.675*	*0.452*	*0.266*	*0.538*
Disadv. Gap [p-value]	[0.26]	[0.153]	[0.2]	[0.528]
C. >= 95% FRL				
SWD Teacher Quality	-0.257	-0.125	-0.164	0.035
Non-SWD Teacher Quality	-0.229	-0.059	-0.124	0.04
TQG	-0.028*	-0.066	-0.04	-0.005
Std. Error	(0.013)	(0.033)	(0.078)	(0.004)
SWD n	13,408	4,988	1,172	15,150
non-SWD n	145,649	52,218	12,808	164,247
between variance	*0.33*	*0.398*	*0.656*	*0.612*
within variance	*0.67*	*0.602*	*0.344*	*0.388*
Disadv. Gap [p-value]	[0.095]	[0.28]	[0.402]	[0.141]

TQG stands for Teacher Quality Gap. GET stands for General Education Teacher. SWD stands for Students with Disabilities. Between/Within variance calculated from Eq. (2) with the inclusion of school fixed-effects. Disadv. Gap represents how similar the FRL bins from the least disadvantaged bin (<70% FRL). Observations are at student-teacher cell level pooled across school years 2014-15 to 2016-17. A school's FRL bin is defined by taking a three year average of the percent of FRL eligible students. Novice is defined by any teacher with fewer than 2 years of experience. VAM, EDST, and Hire are z-scored measures for value-added, evaluation scores, and hiring scores, respectively. *p<0.05, **p<0.01, ***p<0.001

Appendix Table 2D. ELA Teacher Quality Gaps by FRL Bins, Grades 6-8 only				
SWD with GET vs non-SWD				
	(1) VAM	(2) Teach Eval	(3) Hiring Score	(4) Novice
A. <70% FRL				
SWD Teacher Quality	-0.475	0.03	0.115	0.028
Non-SWD Teacher Quality	-0.417	0.225	0.132	0.022
TQG	-0.059*	-0.195**	-0.017	0.006
Std. Error	(0.029)	(0.066)	(0.090)	(0.007)
SWD n	5,125	1,395	525	5,494
non-SWD n	69,405	18,326	4,552	73,220
between variance	*0.648*	*0.533*	*0.69*	*0.77*
within variance	*0.352*	*0.467*	*0.31*	*0.23*
B. 70% - <95% FRL				
SWD Teacher Quality	-0.849	-0.14	0.399	0.057
Non-SWD Teacher Quality	-0.795	0.053	0.249	0.042
TQG	-0.053	-0.193**	0.15	0.015
Std. Error	(0.043)	(0.067)	(0.113)	(0.009)
SWD n	7,892	2,421	1,148	8,515
non-SWD n	80,164	23,615	9,272	85,872
between variance	*0.595*	*0.432*	*0.607*	*0.379*
within variance	*0.405*	*0.568*	*0.393*	*0.621*
Disadv. Gap [p-value]	[0.917]	[0.979]	[0.248]	[0.403]
C. >= 95% FRL				
SWD Teacher Quality	-0.789	-0.052	-0.001	0.071
Non-SWD Teacher Quality	-0.761	-0.01	-0.016	0.057
TQG	-0.028	-0.042	0.016	0.014**
Std. Error	(0.030)	(0.044)	(0.051)	(0.005)
SWD n	14,647	5,713	2,653	16,029
non-SWD n	153,370	53,727	21,014	168,049
between variance	*0.585*	*0.373*	*0.465*	*0.515*
within variance	*0.415*	*0.627*	*0.535*	*0.485*
Disadv. Gap [p-value]	[0.463]	[0.053]	[0.751]	[0.313]

TQG stands for Teacher Quality Gap. GET stands for General Education Teacher. SWD stands for Students with Disabilities. Between/Within variance calculated from Eq. (2) with the inclusion of school fixed-effects. Disadv. Gap represents how similar the FRL bins from the least disadvantaged bin (<70% FRL). Observations are at student-teacher cell level pooled across school years 2014-15 to 2016-17. A school's FRL bin is defined by taking a three year average of the percent of FRL eligible students. Novice is defined by any teacher with fewer than 2 years of experience. VAM, EDST, and Hire are z-scored measures for value-added, evaluation scores, and hiring scores, respectively. *p<0.05, **p<0.01, ***p<0.001

Appendix Table 3. Teacher Quality Gap Variance Decomposition by Subject and FRL Bins

	Math				ELA			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	VAM	Teach Eval	Hiring Score	Novice	VAM	Teach Eval	Hiring Score	Novice
A. <70% FRL								
between variance	*0.260*	*0.378*	*0.427*	*0.041*	*0.439*	*0.393*	*0.510*	*0.449*
within variance	*0.740*	*0.622*	*0.573*	*0.959*	*0.561*	*0.607*	*0.490*	*0.551*
B. 70% - <95% FRL								
between variance	*0.281*	*0.424*	*0.610*	*0.357*	*0.508*	*0.358*	*0.553*	*0.264*
within variance	*0.719*	*0.576*	*0.390*	*0.643*	*0.492*	*0.642*	*0.447*	*0.736*
C. >= 95% FRL								
between variance	*0.240*	*0.298*	*0.453*	*0.393*	*0.451*	*0.317*	*0.390*	*0.448*
within variance	*0.760*	*0.702*	*0.547*	*0.607*	*0.549*	*0.683*	*0.610*	*0.552*

Between/Within variance represent the variance of teacher quality gaps between and within schools found in Table 2. They are calculated from Eq (1) with the addition of school fixed-effects.

Apx Table 4. ELA Teacher Quality Gaps by Disability Type (vs. No Disability) and FRL Bins							
	SWD with GET vs non-SWD				SWD with SET vs non-SWD		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Disability & FRL Group	VAM	Teach Eval	Hiring Score	Novice	Teach Eval	Hiring Score	Novice
A. Specific Learning							
<70% FRL TQG	-0.033	-0.123***	0.055	0	-0.28	-0.604**	0.128**
Std. Error	(0.031)	(0.034)	(0.066)	(0.004)	(0.166)	(0.226)	(0.040)
SWD n	5,131	1,942	630	7,901	535	513	1,931
non-SWD n	121,328	52,275	16,386	218,257	52,275	16,386	218,257
70-<95% FRL TQG	-0.022	-0.103**	0.132	0.006	-0.171	-0.394*	0.116***
Std. Error	(0.043)	(0.038)	(0.086)	(0.005)	(0.090)	(0.158)	(0.022)
SWD n	8,231	3,505	1,032	11,902	1,595	916	5,113
non-SWD n	135,557	66,028	17,872	229,389	66,028	17,872	229,389
Disadv. Gap [p-value]	[0.833]	[0.695]	[0.478]	[0.412]	[0.563]	[0.447]	[0.795]
>=95% FRL TQG	-0.043*	-0.039*	0.031	0.010***	-0.201***	-0.181*	0.124***
Std. Error	(0.022)	(0.019)	(0.038)	(0.003)	(0.048)	(0.071)	(0.012)
SWD n	20,269	10,141	3,283	31,879	6,115	4,664	16,519
non-SWD n	340,481	199,019	54,715	660,545	199,019	54,715	660,545
Disadv. Gap [p-value]	[0.785]	[0.03]	[0.756]	[0.044]	[0.646]	[0.074]	[0.923]
B. Autism							
<70% FRL TQG	0.069	-0.009	0.049	-0.001	-0.065	-0.36	0.118***
Std. Error	(0.042)	(0.039)	(0.075)	(0.004)	(0.088)	(0.182)	(0.023)
SWD n	1,485	690	210	2,905	798	637	2,513
non-SWD n	121,328	52,275	16,386	218,257	52,275	16,386	218,257
70-<95% FRL TQG	0.015	-0.011	0.099	0.005	-0.131	-0.294	0.130***
Std. Error	(0.053)	(0.051)	(0.104)	(0.006)	(0.086)	(0.157)	(0.019)
SWD n	1,165	632	171	2,106	1,250	775	3,450
non-SWD n	135,557	66,028	17,872	229,389	66,028	17,872	229,389
Disadv. Gap							
Disadv. Gap [p-value]	[0.42]	[0.979]	[0.701]	[0.383]	[0.59]	[0.782]	[0.684]
>=95% FRL TQG	0.071	0.03	0.03	0	-0.112*	-0.126*	0.169***
Std. Error	(0.045)	(0.027)	(0.054)	(0.003)	(0.047)	(0.061)	(0.012)
SWD n	2,273	1,435	418	4,613	4,041	3,446	10,191
non-SWD n	340,481	199,019	54,715	660,545	199,019	54,715	660,545
Disadv. Gap [p-value]	[0.974]	[0.402]	[0.831]	[0.852]	[0.635]	[0.223]	[0.052]
C. Speech/Language							
<70% FRL TQG	0.207***	-0.013	0.012	0.001	-0.224	-0.375	0.077*
Std. Error	(0.058)	(0.042)	(0.084)	(0.005)	(0.187)	(0.216)	(0.033)
SWD n	936	814	274	3,134	55	39	147
non-SWD n	121,328	52,275	16,386	218,257	52,275	16,386	218,257
70-<95% FRL TQG	0.308***	0.05	-0.025	-0.003	-0.065	-0.275	0.084***
Std. Error	(0.049)	(0.040)	(0.086)	(0.004)	(0.093)	(0.179)	(0.023)
SWD n	1,242	1,134	233	3,652	145	79	384
non-SWD n	135,557	66,028	17,872	229,389	66,028	17,872	229,389
Disadv. Gap [p-value]	[0.182]	[0.277]	[0.758]	[0.548]	[0.445]	[0.722]	[0.858]
>=95% FRL TQG	0.230***	0.058**	-0.016	-0.008***	-0.148**	-0.086	0.134***
Std. Error	(0.033)	(0.021)	(0.035)	(0.002)	(0.055)	(0.072)	(0.014)
SWD n	3,887	4,296	907	13,749	736	462	1,506
non-SWD n	340,481	199,019	54,715	660,545	199,019	54,715	660,545
Disadv. Gap [p-value]	[0.732]	[0.134]	[0.76]	[0.099]	[0.696]	[0.204]	[0.113]

TQG stands for Teacher Quality Gap. GET stands for General Education Teacher. SET stands for Special Education Teacher. SWD stands for Students with Disabilities. Standard errors clustered at school level. Disadv. Gap represents how similar the Free-Reduced Lunch bins are from the least disadvantaged bin (<70% FRL TQG FRL). Observations are at student-teacher cell level pooled across school years 2014-15 to 2017-18. A school's FRL bin is defined by taking a four year average of the percent of FRL eligible students. Novice is defined by any teacher with fewer than 2 years of experience. VAM (value-added measure), Teacher Eval, and Hiring Score are z-scored measures. *p<0.05, **p<0.01,

Apx Table 5. Math VAM Teacher Quality Gaps by VAM Approach and FRL Bins			
	SWD with GET vs non-SWD		
	(1)	(2)	(3)
	One-step VAM	Two-step VAM, w/o class & grade covs	Two-step VAM
A. <70% FRL			
TQG	-0.136***	-0.203	-0.071***
Std. Error	(0.029)	(0.147)	(0.019)
SWD n	6,090	6,090	6,090
non-SWD n	72,841	72,841	72,841
B. 70% - <95% FRL			
TQG	-0.104***	-0.047*	-0.017
Std. Error	(0.025)	(0.018)	(0.025)
SWD n	8,022	8,022	8,022
non-SWD n	85,772	85,772	85,772
Disadv. Gap [p-value]	[0.401]	[0.294]	[0.086]
C. >= 95% FRL			
TQG	-0.079***	-0.048**	-0.02
Std. Error	(0.013)	(0.017)	(0.011)
SWD n	18,846	18,846	18,846
non-SWD n	209,948	209,948	209,948
Disadv. Gap [p-value]	[0.073]	[0.294]	[0.023]

TQG stands for Teacher Quality Gap. GET stands for General Education Teacher. SWD stands for Students with Disabilities. Standard errors clustered at school level. Disadv. Gap represents how similar the Free-Reduced Lunch bins are from the least disadvantaged bin (<70% FRL). Observations are at student-teacher cell level pooled across school years 2014-15 to 2017-18. A school's FRL bin is defined by taking a four year average of the percent of FRL eligible students. "One-step VAM" directly controls for student demographics in VAM calculations, while "Two-step VAM" residualizes student demographics before calculating VAMs (see Appendix B for more details about VAM construction). *p<0.05, **p<0.01, ***p<0.001

Appendix Table 6. Alternative Math Teacher Evaluation Score Measures by School FRL Bins

	(1) Previous Eval	(2) Resid Previous Eval	(3) Theta (all, previous eval)	(4) Resid Theta (all, previous eval)	(5) Theta (main 3, previous eval)	(6) Resid Theta (main 3, previous eval)	(7) Discussion Techniques and Student Participation (3b2)	(8) Standards-Based Projects, Activities, Assignments (3c1)	(9) Student Feedback (3d3)
A. <70% FRL									
Teacher Quality	0.355	0.012	0.29	0.075	0.284	-0.013	2.795	2.897	2.85
Std. Dev	(0.733)	(0.941)	(0.807)	(0.801)	(0.758)	(0.959)	(0.414)	(0.320)	(0.361)
n	57,150	56,277	56,913	56,361	57,133	56,277	57,133	57,113	57,133
B. 70% - <95% FRL									
Teacher Quality	0.18	0.072	0.084	0.024	0.096	0.019	2.724	2.817	2.778
Std. Dev	(0.829)	(0.960)	(0.904)	(0.908)	(0.897)	(0.992)	(0.463)	(0.413)	(0.448)
n	74,068	72,527	73,826	72,544	74,051	72,527	73,966	74,008	73,989
Disadv. Gap [p-value]	[0.01]	[0.44]	[0.008]	[0.497]	[0.004]	[0.588]	[0.03]	[0.001]	[0.027]
C. >= 95% FRL									
Teacher Quality	0.154	0.035	0.073	0.08	0.08	0.033	2.73	2.812	2.753
Std. Dev	(0.855)	(0.959)	(0.918)	(0.929)	(0.912)	(0.990)	(0.475)	(0.418)	(0.460)
n	226,652	218,235	225,872	219,065	226,498	218,235	226,337	226,460	226,227
Disadv. Gap [p-value]	[0]	[0.733]	[0]	[0.922]	[0]	[0.38]	[0.012]	[0]	[0]

Disadv. Gap represents how similar the FRL bins from the least disadvantaged bin (<70% FRL). Observations are at student-teacher cell level pooled across school years 2014-15 to 2017-18. A school's FRL bin is defined by taking a four year average of the percent of FRL eligible students. *Previous Eval* is the teacher's most recent evaluation score, before the current year. *Resid Previous Eval* represents the residualized teacher evaluation score after accounting for classroom- and school-level demographics. *Theta (all)* is the theta score calculated with the graded response model across all evaluation subcomponents. *Resid Theta (all)* is the residualized theta score after accounting for classroom- and school-level demographics. *Theta (req. 3)* is the theta score calculated with the graded response model, but only with the three required subcomponents. *Resid Theta (req. 3)* is the residualized theta score from the three required subcomponents after accounting for classroom- and school-level demographics. Columns (1)-(6) are standardized by year while Columns (7)-(9) represent raw scores. Sample restricted to observations with valid responses for every outcome.

Appendix Table 7. Math Teacher Quality Gaps by Alternative Evaluation Measures and FRL Bins

	SWD with GET vs non-SWD								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Previous Eval	Resid Previous Eval	Theta (all)	Resid Theta (all)	Theta (req. 3)	Resid Theta (req. 3)	Discussion Techniques and Student Participation (3b2)	Standards-Based Projects, Activities, Assignments (3c1)	Student Feedback (3d3)
A. <70% FRL									
TGQ	-0.027	0.047	-0.052	-0.028	-0.042	0.011	-0.017	-0.01	-0.023
Std. Error	(0.029)	(0.030)	(0.030)	(0.030)	(0.028)	(0.032)	(0.014)	(0.010)	(0.013)
SWD n	3,880	3,880	3,880	3,880	3,880	3,880	3,880	3,880	3,880
non-SWD n	45,269	45,269	45,269	45,269	45,269	45,269	45,269	45,269	45,269
B. 70% - <95% FRL									
TGQ	-0.038	-0.034	-0.027	-0.01	-0.031	-0.018	-0.012	-0.02	-0.01
Std. Error	(0.029)	(0.036)	(0.028)	(0.027)	(0.028)	(0.026)	(0.013)	(0.015)	(0.015)
SWD n	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000
non-SWD n	55,108	55,108	55,108	55,108	55,108	55,108	55,108	55,108	55,108
Disadv. Gap [p-value]	[0.782]	[0.087]	[0.544]	[0.657]	[0.791]	[0.49]	[0.799]	[0.577]	[0.505]
C. >= 95% FRL									
TGQ	-0.023	0.007	-0.031	-0.029	-0.013	0.029	-0.010	-0.007	0.008
Std. Error	(0.016)	(0.015)	(0.017)	(0.017)	(0.017)	(0.016)	(0.010)	(0.006)	(0.006)
SWD n	14,466	14,466	14,466	14,466	14,466	14,466	14,466	14,466	14,466
non-SWD n	166,316	166,316	166,316	166,316	166,316	166,316	166,316	166,316	166,316
Disadv. Gap [p-value]	[0.902]	[0.236]	[0.548]	[0.978]	[0.386]	[0.601]	[0.691]	[0.851]	[0.031]

TQG stands for Teacher Quality Gap. GET stands for General Education Teacher. SWD stands for Students with Disabilities. Disadv. Gap represents how similar the FRL bins from the least disadvantaged bin (<70% FRL). Observations are at student-teacher cell level pooled across school years 2014-15 to 2017-18. A school's FRL bin is defined by taking a four year average of the percent of FRL eligible students. *Previous Eval* is the teacher's most recent evaluation score, before the current year. *Resid Previous Eval* represents the residualized teacher evaluation score after accounting for classroom- and school-level demographics. *Theta (all)* is the theta score calculated with the graded response model across all evaluation subcomponents. *Resid Theta (all)* is the residualized theta score after accounting for classroom- and school-level demographics. *Theta (req. 3)* is the theta score calculated with the graded response model, but only with the three required subcomponents. *Resid Theta (req. 3)* is the residualized theta score from the three required subcomponents after accounting for classroom- and school-level demographics. Columns (1)-(6) are standardized by year while Columns (7)-(9) represent raw scores. Sample restricted to observations with valid responses for every outcome.

Appendix Table 8. SWD with SET vs Non-SWD Teacher Quality Gaps by Subject and FRL Bins						
	Math			ELA		
	(1)	(2)	(3)	(4)	(5)	(6)
	Teach Eval	Hiring Score	Novice	Teach Eval	Hiring Score	Novice
A. <70% FRL						
TQG	-0.203	-0.263	0.086***	-0.170	-0.511**	0.123***
Std. Error	(0.119)	(0.152)	(0.018)	(0.090)	(0.175)	(0.024)
SWD n	1,468	1,089	5,933	1,834	1,531	6,086
non-SWD n	51,285	15,280	207,006	52,275	16,386	218,257
B. 70% - <95% FRL						
TQG	-0.183*	-0.228*	0.103***	-0.090	-0.350*	0.120***
Std. Error	(0.077)	(0.115)	(0.016)	(0.069)	(0.139)	(0.017)
SWD n	4,118	2,639	11,717	4,117	2,497	12,111
non-SWD n	64,067	15,348	225,346	66,028	17,872	229,389
Disadv. Gap [p-value]	[0.885]	[0.853]	[0.478]	[0.479]	[0.47]	[0.91]
C. >= 95% FRL						
TQG	-0.115*	-0.118	0.134***	-0.150***	-0.158**	0.138***
Std. Error	(0.048)	(0.074)	(0.009)	(0.039)	(0.057)	(0.010)
SWD n	13,017	9,334	33,795	13,307	10,379	34,645
non-SWD n	196,697	45,665	653,897	199,019	54,715	660,545
Disadv. Gap [p-value]	[0.493]	[0.392]	[0.019]	[0.84]	[0.055]	[0.562]

TQG stands for Teacher Quality Gap. SET stands for Special Education Teacher. SWD stands for Students with Disabilities. Disadv. Gap represents how similar the Free-Reduced Lunch bins are from the least disadvantaged bin (<70% FRL). Observations are at student-teacher cell level pooled across school years 2014-15 to 2017-18. A school's FRL bin is defined by taking a four year average of the percent of FRL eligible students. Novice is defined by any teacher with fewer than 2 years of experience. VAM (value-added measure), Teacher Eval, and Hiring Score are z-scored measures. *p<0.05, **p<0.01, ***p<0.001

Apx Table 9. Math Teacher Quality Gaps by Disability Type (vs. No Disability) and FRL Bins			
	SWD with SET vs non-SWD		
	(1)	(2)	(3)
Disability & FRL Group	Teach Eval	Hiring Score	Novice
A. Specific Learning			
<70% FRL TQG	-0.408	-0.367	0.058*
Std. Error	(0.288)	(0.197)	(0.025)
SWD n	429	314	2,109
non-SWD n	51,285	15,280	207,006
70-<95% FRL TQG	-0.321**	-0.255	0.100***
Std. Error	(0.096)	(0.161)	(0.024)
SWD n	1,876	1,256	5,481
non-SWD n	64,067	15,348	225,346
Disadv. Gap [p-value]	[0.775]	[0.662]	[0.228]
>=95% FRL TQG	-0.158*	-0.17	0.119***
Std. Error	(0.070)	(0.101)	(0.012)
SWD n	6,435	4,324	17,008
non-SWD n	196,697	45,665	653,897
Disadv. Gap [p-value]	[0.398]	[0.375]	[0.03]
B. Autism			
<70% FRL TQG	-0.068	-0.168	0.100***
Std. Error	(0.094)	(0.173)	(0.022)
SWD n	730	567	2,510
non-SWD n	51,285	15,280	207,006
70-<95% FRL TQG	-0.166	-0.166	0.122***
Std. Error	(0.088)	(0.128)	(0.018)
SWD n	1,260	801	3,455
non-SWD n	64,067	15,348	225,346
Disadv. Gap [p-value]	[0.448]	[0.991]	[0.442]
>=95% FRL TQG	-0.100*	-0.071	0.166***
Std. Error	(0.044)	(0.069)	(0.012)
Std. Error	3,955	3,239	10,075
SWD n	196,697	45,665	653,897
non-SWD n	[0.762]	[0.599]	[0.008]
C. Speech/Language			
<70% FRL TQG	-0.219	-0.252	0.057
Std. Error	(0.161)	(0.275)	(0.030)
SWD n	50	34	154
non-SWD n	51,285	15,280	207,006
70-<95% FRL TQG	-0.16	-0.205	0.059**
Std. Error	(0.097)	(0.144)	(0.019)
SWD n	155	90	396
non-SWD n	64,067	15,348	225,346
Disadv. Gap [p-value]	[0.752]	[0.88]	[0.973]
>=95% FRL TQG	-0.139*	-0.028	0.136***
Std. Error	(0.058)	(0.082)	(0.014)
SWD n	724	437	1,498
non-SWD n	196,697	45,665	653,897
Disadv. Gap [p-value]	[0.638]	[0.434]	[0.018]

TQG stands for Teacher Quality Gap. GET stands for General Education Teacher. SWD stands for Students with Disabilities. Standard errors clustered at school level. Disadv. Gap represents how similar the Free-Reduced Lunch bins are from the least disadvantaged bin (<70% FRL TQG FRL). Observations are at student-teacher cell level pooled across school years 2014-15 to 2017-18. A school's FRL bin is defined by taking a four year average of the percent of FRL eligible students. Novice is defined by any teacher with fewer than 2 years of experience. VAM (value-added measure), Teacher Eval, and Hiring Score are z-scored measures. *p<0.05, **p<0.01, ***p<0.001

Appendix A. Teacher Assignment

(via Los Angeles Unified School District Human Resources: Staff Relations Handbook)

This information is intended to provide guidance to Principals so that they can comply with the LAUSD/UTLA (United Teachers Los Angeles) Collective Bargaining Agreement and assure that teacher assignments best meet students' instructional needs and priorities.

Elementary School Assignments

1. In elementary schools, the LAUSD/UTLA Collective Bargaining Agreement (**CBA**) Article IX-A, Section 2.0 c (1) (ii) provides that the site administrator shall assign permanent teachers to **track** or **grade level** opening on the basis of seniority. Appropriate credential should be considered for Special Education assignments.
2. The Collective Bargaining Agreement does not provide teachers the right to select specific instructional programs, student performance levels or instructional clustering of students.
3. Principals can use preference forms (District's preferred method that will support an effective instructional program) or locally determined method to receive teachers' requests for assignments.
4. The site administrator **can and should make exceptions to the CBA provision** if he or she reasonably determines that the specific assignment is not in the best interest of the educational program.
5. Teachers with the specified credentials and required qualifications ("qualified") may request assignment to their grade level using a teacher preference form or other locally determined method. Submission of this form shall serve as a request for the assignment.
6. Administrators should review credentials, specific training, authorizations, performance indicators (i.e. pre/post assessment data, ELL reclassification data, DIBELS) teacher status (Probationary 1 and 2) and evaluation/conduct records to inform their decision to assign a teacher to a specific class.

Secondary School Assignments

1. In secondary schools, Article IX-A, Section 2.0 d only provides teachers the right to a **department** selection on the basis of recent experience/seniority.
2. **Principals retain the authority** to assign teachers to particular classes and sections within a department.
3. Secondary principals must understand that the CBA does not confer the right for teachers to select either classes or "lines" on the master schedule.

4. Principals should take in consideration the best interest of the instructional program including specific training, authorizations, performance indicators (i.e. core subject end of the year assessments data, ELL reclassification data, pre/post assessment data) teacher status (Probationary 1 and 2) and evaluation/conduct records to inform their decision to assign a teacher to a specific class.

5. Principals **can and should use objective data** as described above to assign teachers to classes.

6. Classes within a department shall be distributed by the Principal (or designee) in consultation with the **elected department chair**.

Appendix B. Two-step Average Residual and One-Year VAM Calculations

Two-step Average Residual VAM Calculations

We begin by using the following equation to create a residualized test score for student *i* in year *t*:

(1) $$Ach_{ijst} = \delta_1 SameAch_{ijst-1} + \delta_2 SameAch^2_{ijst-1} + \delta_3 SameAch^3_{ijst-1} + \delta_4 OtherAch_{ijst-1} + \delta_5 OtherAch^2_{ijst-1} + \delta_6 OherAch^3_{ijst-1} + \boldsymbol{X}_{ijst}\boldsymbol{\theta} + \boldsymbol{T}_{jt}\boldsymbol{\Omega} + \varepsilon_{ijst}$$

where Ach_{ijst} is either math or ELA achievement, standardized within test and year, for student *i* with teacher *j* in school *s* and year *t*, $SameAch_{ijst-1}$ is the student's prior year score in the same subject, which enter as a cubic polynomial, $OtherAch_{ijst-1}$ is the student's prior year score in the other subject, which also enters as a cubic polynomial. In other words, if we are looking at math achievement, $SameAch_{ijst-1}$ would represent the student's prior math test score and $OtherAch_{ijst-1}$ would represent the student's prior ELA test score. X is a vector of student-, classroom- and grade-level demographics, and $\boldsymbol{T}_{jt}$ is a vector of teacher fixed effects. Specifically, X contains information about race, gender, free/reduced lunch status, English Language Learner status, student with disability status, and testing accommodation flags for technology (i.e. text-to-speech software), setting (i.e. small group setting), time (i.e. extended time), and format (i.e. streamlined version of text). Consequently, the residualized test scores are calculated in the following equation:

$$(2)\quad Ach^{*}_{ijst} = SameAch_{ijst} - \delta_1 SameAch_{ijst-1} - \delta_2 SameAch^{2}_{ijst-1} - \delta_3 SameAch^{3}_{ijst-1} - \delta_4 OtherAch_{ijst-1} - \delta_5 OtherAch^{2}_{ijst-1} - \delta_6 OherAch^{3}_{ijst-1} - \boldsymbol{X}_{ijst}\boldsymbol{\theta} = \boldsymbol{T}_{jt}\boldsymbol{\Omega} + \varepsilon_{ijst}$$

Next, students' residual scores in time *t* are averaged to create teacher value-added for teacher *j*, $\overline{A}_{jt}$. Residual average scores from prior years are used to calculate the best linear predictor of $\overline{A}_{jt}$ for teacher *j* in year *t* and forecasting coefficients, ψ, that minimizes the mean-squared error of the test score forecasts are selected:

$$(3)\qquad \psi = \underset{\{\psi_1,\ldots,\psi_{t-1}\}}{\operatorname{argmin}} \sum_j \left(\overline{A}_{jt} - \sum_{s=1}^{t-1} \psi_s \overline{A}_{js}\right)^2$$

Finally, estimates of ψ from any year outside of *t* are used to calculate value-added for teacher *j* in year *t*.

<u>One-step VAM Calculations</u>

We estimate VAMs using the following model for each school year from 2012-2013 to 2016-17, and separately for each subject and level (elementary and secondary):

$$(4)\qquad Ach_{ijst} = \beta_1 Ach^{math}_{ijst-1} + \beta_2 Ach^{ela}_{ijst-1} + \boldsymbol{X}_{ijst}\boldsymbol{\theta} + \boldsymbol{T}_{jt}\boldsymbol{\Omega} + \varepsilon_{ijst}$$

where Ach is either math or ELA achievement, standardized within test and year, for student *i* with teacher *j* in school *s* and year *t*. We control for students' achievement in the prior year in both math and ELA, since the inclusion of the second subject is helpful in mitigating bias due to sorting (Chetty, Friedman, & Rockoff, 2014) and to attenuate measurement error (Lockwood & McCaffrey, 2014). **X** is a vector of student demographic characteristics, including indicators of student race, gender, free- or reduced-price lunch eligibility, grade level, and English learner status. Johnson & Semmelroth (2013) argue that each disability type requires different teaching methods so we also include disability type indicators (Autism, Specific

Learning Disability, Speech/Language Impairment, and Other Disability) into the vector **X** in order to measure a teacher's overall effectiveness in achievement growth, as opposed to effectiveness towards specific disability types. The final component of **X** is an indicator which describes whether a student has testing accommodations. Specifically, we include four types of testing accommodation flags: technology (i.e. text-to-speech software), setting (i.e. small group setting), time (i.e. extended time), and format (i.e. streamlined version of text). Jones, Buzick, & Turkin (2013) argue that testing accommodations influence a student's test score in an ambiguous manner. If left out, accommodations could introduce measurement error into the VAM scores. Teachers' VAMs are estimated by the coefficients on a set of teacher fixed effects (**T**), and ε is an error term. This specification was chosen based on a detailed review of the current best practices in VAM modeling, summarized in Koedel, Mihaly, and Rockoff (2015).

We use teacher-year models instead of models that pool data over time because we are interested in teachers' effectiveness in the specific year they taught each student, not teachers' average VAM over time. As a robustness check, we also examined how our value-added estimates varied across other commonly used alternative specifications (Herrmann, Walsh, Isenberg, & Resch, 2013; Koedel et al., 2015). Our preferred model is consistently highly correlated with these alternative specifications (0.94 or above).

Since previous standardized test scores are required to calculate VAMs, we are only able to calculate VAMs for fourth through eighth grade teachers. 45% of our initial sample has a valid math VAM, or 395,426 student-year observations. Following Goldhaber et al. (2015), in our teacher quality gap models (equations 1 and 2) we use each teacher's VAM estimate from the prior school year so that students' current test scores are not taken into consideration.

Our estimates constrain the teacher fixed effect estimates to sum to zero so that teachers are compared to the average score instead of an omitted teacher. Each observation is weighted by the share of trimesters or semesters in the year during which the student-teacher link was observed.

Appendix C. LAUSD's Multiple Measures Teacher Selection Process (from Bruno and Strunk, 2019)

"Since SY2014-2015, teacher applications are processed through the following sequence. Applications are first checked for completeness and if they meet the minimum criteria. Applicants are disqualified if their application is incomplete, if their credentials are inadequate, or if there are no vacancies for particular positions. For those who pass the first round, LAUSD reaches out for profession references and ask candidates to complete an online written assessment, asking teachers to describe how they would respond to a series of different vignettes. Applicants who do not provide professional references, receive an "ineffective" rating from their references, or score lower than 11 points on their written sample are eliminated. Remaining applicants are invited to the district office for a structured formal interview and to provide sample lesson demonstrations, which are scored by HR specialists. Initial applications are scored (based on undergraduate grade point average, subject matter preparation, and background) and added to the overall applicant score from the interview, professional references, sample less, and writing sample. Applicants who receive at least 80 points and meet all the minimum required scores for each component (detailed in the Table below) are placed on an eligibility list. However, there are two possible exceptions. One, school principals can request that a specific candidate receive an exception to a score requirement (thought they still have to go through the application process). Two, candidates who fail to meet the minimum score requirement for one component, or only do not meet the 80-point requirement, are resubmitted to an HR specialist panel for a blind review. If the panel agrees that the candidate is high-quality, the candidate is added to the eligibility list. Schools draw from the eligibility list to hire for their vacancies and have flexibility in how they wish to interview/screen these candidates. School administrators never obtain the hiring scores, they only know if the candidate is eligible for hire."

Eligibility Criteria for Prospective Teachers in LAUSD

Criterion	Description	Minimum Points Possible	Maximum Points Possible	Minimum Passing Score
Interview	*Structured, conducted by one HR specialist.*	0	25	20
Professional References	*Collected from student teaching or other past professional experience.*	0	20	16
Sample Lesson	*Delivered to and evaluated by two HR specialists.*	0	15	11
Writing Sample	*Timed (45 minutes) responses to hypothetical student-related scenarios.*	1	15	11
GPA	*Scored based on verified undergraduate GPA.*	1	10	N/A
Subject Matter	*Based on subject-matter licensure test scores or, if waived, GPA score.*	8	10	N/A
Background	*For any of: certain prior LAUSD (non-teaching) experience, prior leadership (e.g., military experience), possession of a graduate degree, or Teach for America experience.*	0	2	N/A
Preparation	*For any of: attendance at school highly-ranked by U.S. News & World Report, evidence of prior teaching effectiveness (e.g., student achievement data), or major in credential subject field or, if multi-subject, core academic subject/liberal arts.*	0	3	N/A
Overall		10	100	80

Note. Points are awarded in accordance with criterion-specific rubrics aligned to district goals (e.g., employee evaluation criteria).
Applicants may be placed on the eligibility list despite scoring below the minimum passing score at the request of a school administrator or upon a review of application materials by human resources staff. (Bruno and Strunk 2019)