Measuring Teacher Effectiveness for Gifted and Talented and High-Testing Students

Seth Walker
Michigan State University

Scott Imberman
Michigan State University, CES-ifo, & NBER

Katharine Strunk
University of Pennsylvania

April 29, 2025

Abstract

This paper investigates the differential effectiveness of teachers across student populations, focusing on gifted and talented (GT), high-testing, and low-testing students. Using data from Los Angeles, we show that there is considerable variation in teachers' relative effectiveness (as proxied by value-added measures of teachers' contributions to student achievement growth) across these subgroups. GT students are more likely to have teachers with higher VAMs across all subgroup measures, and, to a modest extent, are more likely to be matched with teachers who have a relative advantage in teaching GT students.

Keywords: Value-added, gifted and talented, teacher quality

Seth Walker: Department of Economics, Michigan State University. 486 W Circle Dr. East Lansing, MI 48824. walke893@msu.edu

Scott Imberman: Department of Economics, Michigan State University. 486 W Circle Dr. East Lansing, MI 48824. imberman@msu.edu
Katharine Strunk: Graduate School of Education, 3700 Walnut Street Philadelphia, PA 19104. gsedean@gse.upenn.edu

Acknowledgements: This work was funded by the Institute for Education Sciences, US Department of education through grant R305C200012 to University of Connecticut. We thank the Los Angeles Unified School District for providing the data that made this research possible. The opinions expressed are those of the authors and do not represent the views of the Institute, the US Department of Education, or the Los Angeles Unified School District.

1 Introduction

The identification of high-quality teachers and their practices holds significant interest for a variety of stakeholders. School and district administrators can strategically place students with effective teachers or use their methods to guide others; students and parents value access to exceptional teachers; and teachers themselves deserve appreciation and recognition for their contributions. However, most measures of teacher quality rely on aggregate metrics that average performance across all students in a class, potentially masking differential effectiveness across student subgroups. Understanding variations in teacher effectiveness for different student groups is crucial for stakeholders striving to achieve desired educational outcomes. For instance, a principal or parent seeking to pair a child struggling in math with the most suitable teacher would prioritize a teacher who excels at fostering growth in struggling students, even if their overall quality measure does not reflect this specific capability.

Teachers may exhibit varying levels of effectiveness for different student subgroups, including English Learners (ELs), students with disabilities (SWDs), and gifted and talented (GT) students, among others. Prior research using achievement-based value-added measures (VAMs) has shown that teachers differ in their effectiveness for ELs (Loeb et al., 2014; Master et al., 2016) and SWDs (Wood et al., 2024) relative to students not in these categories.

Moreover, in the case of SWDs, Wood et al. (2024) demonstrates that teachers are not always matched with the students for whom they are most effective. However, there is limited research regarding whether and what kinds of teachers are more effective with gifted or high test scoring students compared to low test scoring students (henceforth, “high-testing” and “low-testing”).

This gap in the literature is significant given that nearly 2 million students across the United States were designated as GT in the 2017-2018 school year.[1] Despite this, inadequate funding (Kettler et al., 2015; Rinn et al., 2020) and concerns over exacerbating social inequities (Dai, 2013) — especially the underrepresentation of students of color (Peters and Engerrand, 2016; Grissom and Redding, 2016; Olszewski-Kubilius and Steenbergen-Hu, 2017) — result in many GT students spending the majority of their day in general classroom settings with teachers who are not trained to meet their unique academic needs (Bangel et al., 2010; NAGC, 2014b; Mathijssen et al., 2021).

This study leverages data from the Los Angeles Unified School District (LAUSD) to examine whether teachers differ in their effectiveness across subgroups, with a particular focus on GT students. Using subgroup-specific value-added measures of teachers' contributions to student achievement on standardized tests (henceforth, “achievement value-added”), we estimate each teacher's effectiveness for GT, high-testing, and low-testing students separately. We then investigate whether GT students are more likely to be as-

[1]Source: U.S. Department of Education Office of Civil Rights

signed to teachers with higher overall VAMs, as well as whether they are more likely to be matched with teachers who have a relative advantage in teaching GT students compared to low-testing peers.

Our findings reveal meaningful differences in teachers' effectiveness across student subgroups in both math and ELA. However, these differences are not well explained by observable teacher characteristics, which are generally unrelated to subgroup VAMs. GT students are more likely to be assigned to teachers with higher VAMs across all three subgroup measures, and we find modest evidence that they are more likely to be matched with teachers who have a relative advantage in teaching GT students. This pattern is most pronounced in ELA for students classified as highly gifted, who are 4–6 percentage points more likely to be assigned to teachers with a relative advantage in teaching GT students — approximately four times larger than the corresponding estimate for all GT students. Taken together, these results suggest that assignment patterns partly reflect subgroup-specific instructional strengths.

2 Brief Review of the Literature

The relationship between teacher characteristics and their effectiveness with high-testing students, including those identified as gifted, has been a persistent concern in educational research. One major challenge is that few teacher characteristics reliably predict effectiveness for high-testing students.

This is likely because many educators receive insufficient training in gifted education, which directly impacts their preparedness and confidence in teaching advanced learners (Gallagher, 2015). In fact, many teachers lack resources and the necessary professional development to understand and implement strategies that are beneficial for GT students (Sahin and Levent, 2015; Barbier et al., 2023). This lack of training often leads to low confidence in instructing GT students and a reluctance to employ differentiated instructional techniques (Gross, 1994; Dixon et al., 2014; Fraser-Seeto et al., 2015; Matheis et al., 2017). Instead of differentiating instruction, teachers often teach to the level of either the average student or struggling students (Hansen and Feldhusen, 1994), leaving GT students unchallenged and with few opportunities for growth (Berman et al., 2012; NAGC, 2014a).

The impact of inadequate training is compounded by misconceptions and attitudes that teachers hold about GT students. Some teachers view special provisions for GT students as unnecessary or even unfair (Berman et al., 2012), while others stereotype GT students as "socially noncompliant" or arrogant (Geake and Gross, 2008). Such misconceptions are particularly prevalent among more experienced teachers, who may feel less inclined to meet the needs of GT students (Perković Krijan and Borić, 2014). This is concerning given that a teacher's attitude has been linked to their instructional quality and effectiveness (David, 2011; Troxclair, 2013). Consequently, teachers harboring these misconceptions may inadvertently undermine the academic growth of GT students.

Moreover, research into the attributes of effective teachers for GT students often highlights qualitative traits unrelated to formal qualifications or certifications. Surveys of GT students reveal a preference for teachers who demonstrate high intelligence (Bishop, 1968), effective classroom management (Mills, 2003), and, most importantly, positive personal attributes such as enthusiasm, flexibility, and the ability to motivate (Maddux et al., 1985, Eilam and Vidergor, 2011; Khalil and Accariya, 2016). Likewise, studies identifying exemplary teachers for the gifted emphasize these qualities, with additional findings suggesting that these characteristics are more predictive of teacher effectiveness than factors like gender, race, or years of experience (Chan, 2001; Woods, 2004; Tischler and Vialle, 2009; Stephens, 2009; David, 2011). Although Chessman (2010) and Lassig (2015) found some evidence that gender may play a role in teacher effectiveness for GT students, the broader literature lacks consensus on this relationship.

While prior research has explored the effectiveness of teachers for English Learners and students with disabilities, there remains a significant gap in understanding how teacher effectiveness varies for high-testing students, including those designated as GT. This gap is particularly concerning given the unique academic needs of GT students and the potential for mismatches in teacher-student assignments that fail to fully leverage a teacher's strengths.

Although limited, recent research has begun to explore dimensions of heterogeneity in teacher effectiveness that relate to, but do not directly address, the assignment of high-testing or GT students to teachers. For example, Biasi et al. (2021) examine heterogeneity in teacher value-added for low- and high-achieving students, finding that teacher effectiveness is positively correlated across groups and that more effective teachers are generally less willing to teach in lower-achieving districts — except when they have a comparative advantage with low-achieving students. However, the study focuses on district-level sorting rather than within-school student–teacher assignments and does not examine whether students are matched to teachers based on subgroup-specific strengths. Moreover, beyond teaching experience, the paper does not explore which observable teacher characteristics are associated with differential effectiveness across student performance levels.

Similarly, Ahn et al. (2025) develop a multidimensional value-added framework that accounts for teacher–student match effects across several dimensions, including prior achievement. They find that most teachers are not equally effective for all students, and that the difference in outcomes between well- and poorly-matched teachers is especially large for low-achieving students. While their approach highlights the significance of heterogeneity in teacher effectiveness and the value of matching students with teachers who are particularly effective for them, it does not focus specifically on gifted students or examine whether schools actively match high-testing students to teachers with a demonstrated comparative advantage.

This paper contributes to the existing literature by examining the differential effectiveness of teachers for GT, high-testing, and low-testing students within LAUSD. Specifically, we use teacher VAMs to assess whether GT and high-testing students are systematically assigned to teachers who are particularly effective at promoting their academic growth. Our analysis extends the understanding of how teacher assignments are made for GT students and whether these assignments align with teachers' demonstrated strengths. In doing so, we complement recent work on heterogeneity in teacher effectiveness by showing that these differences extend to GT students and that assignment patterns partially reflect this variation.

3 Data

3.1 Background and Context

This paper utilizes student- and teacher-level administrative data provided by LAUSD's Office of Data and Accountability and the Division of Human Resources, covering the 2007-08 to 2017-18 academic years. The student-level data include annual observations of student race/ethnicity, gender, free or reduced-price lunch (FRL) eligibility, English Learner (EL) status, GT indicators, and math and English Language Arts (ELA) test scores. We normalize test scores to have a zero mean and standard deviation (SD) of one for each subject-grade-year combination. We limit analysis to students in grades 3 through 8, as these are the only grades for which LAUSD collects test scores. The teacher-level data include information on race/ethnicity, gender, years of experience, and education.

As one of the largest school districts in the United States, LAUSD offers an exceptional context for studying teacher effectiveness across a diverse student body. In particular, with nearly one in six students designated as GT, the district provides an ideal setting for examining teacher effectiveness for academically talented students.

3.2 Generating VAMs

Our primary goal is to explore differences in teacher effectiveness across student subgroups — specifically, whether teachers are differentially effective for GT, high-testing, and low-testing students. To estimate teacher effectiveness for each group, we restrict the sample to teachers who have taught at least 15 students from each subgroup between 2007–08 and 2017–18. This ensures that all value-added measures are based on teachers with meaningful exposure to each student type. We allow for overlap between groups, as students may be both GT and high-testing (or GT and low-testing) in a given year.

We define high-testing students as those whose prior-year test score in a given subject (math or ELA) falls in the top quartile of the grade-year distribution; low-testing students fall in the bottom quartile. To focus cleanly on group-specific effectiveness, we restrict our estimation to students in the high-testing, low-testing, or GT categories and exclude students whose prior scores fall in the middle two quartiles. Given these restrictions, we observe 4,399 unique teachers and 532,491 unique students.

This approach allows us to discern which teachers are most effective for GT and high-testing students, while avoiding arbitrary cutoffs that rely on small differences between academically similar students. For example, the distinction between students just above and below the 75^{th} percentile is often minimal, whereas the contrast between students in the top and bottom quartiles is substantial. Excluding the middle 50% of the distribution enables clearer estimation of whether teacher effectiveness varies across the performance spectrum and between formally identified GT students and other high-achievers.

Panel A of Table 1 presents summary statistics for all LAUSD students, categorized into GT, non-GT high-testing, and non-GT non-high-testing groups. While not identical, high-testing non-GT students share similarities with GT students. In comparison to non-high-testing students, both GT and high-testing non-GT students have significantly lower levels of FRL eligibility and EL or SWD classifications. Additionally, these groups have a higher proportion of Asian and White students. As expected, both GT and high-testing non-GT students also have significantly higher standardized test scores than their non-GT non-high-testing counterparts. However, the scores of GT students are nearly half a SD higher than non-GT high-testing students.

These patterns highlight that, although GT and high-testing students share many characteristics, important differences remain. In addition to having higher average test scores, GT students may be perceived differently by teachers or subject to different expectations as a result of their classification. Although restricting the sample to teachers with sufficient experience across all three groups reduces the analytic sample, it allows for more meaningful comparisons of teacher effectiveness across student types and strengthens the precision and interpretability of our subgroup-specific estimates.

Panel B of Table 1 provides summary statistics for GT and high-testing non-GT students in our analytical sample. The characteristics of these students in the analytical sample closely align with those of their broader population within LAUSD, reflecting similar socioeconomic status, demographic profiles, and academic performance.

3.3 GT Classification

LAUSD classifies GT students into categories based on standardized tests, district-approved assessments, course grades, and demonstrated abilities in non-testing subjects. Table 2 details the specific criteria, evaluation and identification procedures, and the applicable grade levels for each classification, as outlined on the LAUSD's Gifted/Talented Programs website (LAUSD, 2024). Once identified as gifted, a student maintains their classification until graduation.

GT classifications in LAUSD encompass three broad areas: 1) intellectual development, 2) academic achievement in mathematics and/or ELA, and 3) talent in non-testing subjects. Students above the 95^{th} percentile in intellectual development fall into the "Intellectual Ability" (IA) classification, while those in the 99.5^{th} percentile or higher are designated as "Highly-Gifted" (HG). In terms of academic achievement, students who demonstrate exceptional performance on ELA and math assessments for two consecutive years are classified under "High Achievement" (HA). Those who maintain three consecutive years of exceptional performance in either subject receive a "Math" or "ELA" gifted classification. Students exhibiting talent in visual and/or performing arts through an audition or portfolio receive an "Arts" gifted classification. Lastly, students can attain GT status via demonstrating creative or leadership abilities.[2] Students are assigned a single GT classification.

While GT identification in LAUSD is formal process, it does not result in structural separation at the school or classroom level. Figure 1 shows the distribution of GT students by school-year and class-year. The vast majority of schools enroll relatively small shares of GT students, with few schools where GT students make up even 40% of the population. Similarly, GT students are widely dispersed across classrooms: in most class-year settings, fewer than

[2] Fewer than 100 students are in the Creative or Leadership classification in our analytical sample. These students remain in our analysis for all GT students, but we do not include separate estimates for either subgroup.

10% of students are classified as GT, and there are virtually no classrooms composed exclusively of GT students. These patterns show that GT students are overwhelmingly integrated into general education environments, rather than assigned to dedicated GT schools or tracked into GT-only classes.

Table 3 presents descriptive statistics for students and teachers in our analytical sample, with Panel A focusing on students and Panel B on teachers. First, Panel A offers a detailed breakdown of students by GT status, then across the various GT classifications. For students in our sample, 76.7% of 3^{rd} to 8^{th} grade students are non-GT, with the remaining 23.3% identified as GT.[3] GT students demonstrate significantly higher academic performance in both math and ELA, as indicated by test scores over a full standard deviation higher than those of non-GT students. Beyond academic performance, there are distinct demographic differences: GT students are less likely qualify for FRL or be designated as an EL or SWD. They are also more likely to be Asian or White compared to their non-GT counterparts.

Subcategories within the GT classification reveal additional nuances. For example, female students are overrepresented in the ELA (69.5%) and Arts (73.9%) categories, while male students are overrepresented in Math (57.7%) and HG classifications (59.3%). Asian and White students are particularly concentrated in the HG classification, accounting for 20.0% and 32.2% of this

[3] Note that identification does not necessarily mean the student is receiving GT services. It is unclear how this proportion compares to the rest of the country or California as a whole as data on GT rates by grade level are not available. Including all grades, nationally 7% of students are enrolled in a GT program, with a slightly higher 8% in California as of 2013 (U.S. Department of Education, 2017).

group, respectively. On the other hand, GT Math students are predominantly Latino/a (74.2%) and are the most likely to be eligible for FRL (77.8%).

Panel B shifts the focus to the teachers of GT and non-GT students, revealing only slight differences between these two groups. Teachers of GT and non-GT students have similar educational backgrounds, with approximately 40% holding advanced degrees, and comparable levels of teaching experience. However, GT students are marginally more likely to have teachers with National Board Certification than non-GT students (6.6% vs 4.9%). There are also modest differences in the demographic profiles of these teachers: a slightly higher proportion of teachers of GT students are female (69.1% vs. 66.1%) and white (48.4% vs. 43.6%), while a somewhat smaller proportion are Latino/a (25.8% vs. 30.4%).

3.4 VAM Calculations

We follow the methodology of Wood et al. (2024) and calculate separate VAMs for math and ELA test scores for each teacher to measure their contribution to student achievement growth for GT, high- and low-testing students using a leave-year-out estimation model based on Chetty et al. (2014). This approach reduces measurement error. Specifically, we estimate the following model separately by student achievement status:

$$\begin{aligned} Ach_{ijt}^{subject} =& f(Ach_{ijt-1}, Ach_{ijt-2}) + X_{ijt}\Gamma + C_{jt}\Psi + S_{st}\Phi + \\ & T_j\Omega + \theta_t + \epsilon_{ijt}, \end{aligned} \tag{1}$$

where $Ach_{ijt}^{subject}$ represents the standardized (within year) math or ELA achievement score for student i with teacher j in year t. We include a flexible functional form of lagged achievement, $f(Ach_{ijt-1}, Ach_{ijt-2})$, which consists of linear, squared, and cubic terms for each student's math and ELA scores in the previous two years. This flexible specification helps address nonlinearity in the relationship between past and current achievement, while further mitigating potential bias due to sorting on unobserved ability (Lockwood and McCaffrey, 2014).

The vector X_{ijt} controls for student demographics that may influence test scores outside of a teacher's control, including race, gender, FRL eligibility, EL status, SWD status, and grade level. We also control for classroom composition (C_{jt}), including the proportion of students in the class who are FRL, EL, SWD, and GT, as well as class-level means of prior achievement. We also include school-level controls (S_{st}) for the same set of variables used in the classroom composition vector.

Teacher VAMs are estimated by the coefficients on a set of teacher fixed effects (T_j), which are normalized at the teacher-year level, which provide a subgroup-specific VAM estimate for each teacher-year. θ_t are year fixed effects and ϵ_{ijt} is a normally distributed error term. Standard errors are adjusted for heteroskedasticity.

To ensure VAMs are based on meaningful within-year variation, we require that teachers have at least one student in each relevant subgroup in a given year in order to receive a VAM estimate for that group-year. While some teacher-year estimates are based on relatively small samples, the typical group-year contains a non-trivial number of observations: the median number of students per teacher per group is 11, 8, and 6 for high-testing, low-testing, and GT, respectively.

4 RQ1 : How prevalent are teachers who demonstrate a relative advantage in teaching high-testing or GT students compared to low-testing students, and vice versa?

We begin by examining differences in teachers' relative ranks across subgroup-specific VAMs — specifically, how their effectiveness with high-testing and GT students compares to their effectiveness with low-testing students. As VAMs are estimated separately for each teacher and subgroup in each year, the resulting scores are ordinally but not cardinally comparable. That is, we can assess whether a teacher ranks higher or lower in the VAM distribution for one group relative to another, but we cannot compare magnitudes across distributions, as each is standardized relative to the mean in its group.

Following Wood et al. (2024), we define a teacher's relative rank for each VAM (math/ELA by GT/high-/low-testing) as their percentile within the VAM distribution for a given year. We then compute the difference in VAM percentiles (DVAM). For GT versus low-testing students, the DVAM is defined as:

$$DVAM_{jt}^{subject} = Ptile_{jt}(VAM_{Low}^{subject}) - Ptile_{jt}(VAM_{GT}^{subject}) \tag{2}$$

where $Ptile(.)$ converts its argument into a percentile. Using Equation (2), we define a teacher as having a relative advantage in instructing high-testing students if their VAM percentile in the GT student distribution is greater than their VAM percentile in the low-testing student distribution, thus having a negative DVAM. We construct an analogous measure comparing teachers' VAM percentiles for low-testing and high-testing students, allowing us to assess relative advantage in teaching high-testing students.

Figure 2 illustrates this concept by plotting each teacher's VAM percentile ranks for different student subgroups. Panels A and B show the relationship between high-testing and low-testing VAM percentiles in math and ELA, respectively, while Panels C and D present the same comparison for GT and low-testing students. If teachers were equally effective across subgroups, the points would cluster tightly along the 45-degree line. Although a positive relationship exists between the two VAMs in each panel, as indicated by the dotted fitted lines, not all teachers who rank highly in one VAM also rank highly in the other. Teachers above the 45-degree line are in a higher percentile for GT (high-testing) VAMs than low-testing VAMs, indicating a relative advantage in teaching GT (high-testing) students. Conversely, teachers below the 45-degree line have a relative advantage in teaching low-testing students. The distance from the 45-degree line reflects the magnitude of a teacher's comparative advantage.

Although we observe differences in teacher rankings across subgroups, a potential concern with this construct is that VAM estimates are noisy (Hanushek and Rivkin, 2010), meaning that part of the observed within-teacher variation between types of VAMs is due to random error. In particular, measurement error in VAMs may lead to misclassification in our relative advantage indicators. To assess the extent of this error, we calculate the consistency of DVAM classifications across years for each teacher by computing the proportion of years in which a teacher remains on the same side of the DVAM threshold (i.e., consistently classified as having a relative advantage in teaching high- or low-testing students). For math, the average consistency rate is 80.9%, and for ELA it is 83.2%, suggesting that classification error is present but not extreme, and that any resulting attenuation bias is likely to be modest.

As a further check, we construct a set of VAMs based on randomly assigned student subgroups. Specifically, we randomize each student's previous achievement quartile and GT status and then calculate separate VAMs by group for each teacher. Since these groups are randomly assigned, any difference in a teacher's randomized VAMs across these groups would be attributable to random noise. We repeat this process 100 times and construct a randomized DVAM for each subject and randomization, which we then pool together to create a random DVAM distribution. This provides a baseline distribution for comparison with each original DVAM distribution.

This randomization should result in similar subgroup VAMs as any differences would only reflect random error. If the original DVAMs contain information beyond random error, the variation in the randomized DVAM distribution would be smaller than that of the original distribution. In turn, this would support the notion that some teachers possess a relative advantage in teaching certain types of students.

Figure 3 displays the distribution of DVAMs for both actual and randomized subgroup assignments across four panels. Panels A and B show the distribution of DVAMs comparing high-testing and low-testing students in math and ELA, respectively. Panels C and D present the analogous distributions comparing GT and low-testing students. In each case, the solid line represents the actual DVAM distribution, while the dashed line depicts the randomized counterpart.

Across all panels, the original DVAM distributions are visibly more dispersed, with heavier tails and flatter peaks than those generated under random assignment. This indicates that a greater share of teachers display substantial comparative advantage — whether for high-testing, low-testing, or GT students — than would be expected due to random noise alone. We formally test for distributional differences using a two-sample Kolmogorov–Smirnov test, which rejects the null hypothesis that the actual and randomized DVAMs are drawn from the same distribution in all four comparisons ($p < 0.001$). These findings support the conclusion that subgroup-specific teacher effectiveness reflects true variation rather than estimation error.

Table 4 provides further detail on the distribution of subgroup-specific VAM quintiles across teacher-year observations. Panels A and B display the joint distribution of high-testing and low-testing VAM quintiles for math and ELA, respectively, while Panels C and D present the same distributions for GT and low-testing VAMs. Each cell represents the percentage of observations falling into a given combination of quintiles. The bolded diagonal entries indicate observations that fall within the same quintile for both groups (e.g., in the third quintile for VAMs for both high- and low-testing students). Values below the diagonal represent teachers in higher GT (high-testing) VAM quintiles compared to low-testing VAM quintiles, indicating they are relatively more effective in teaching GT (high-testing) students. Conversely, values above the diagonal represent teachers who are relatively more effective at teaching low-testing students.

Across all panels, a large portion of teachers are either minimally or maximally effective across both groups. For example, 10.30% of math teachers fall in the bottom quintile for both high-testing and low-testing students, while 11.28% fall in the top quintile for both. A similar pattern emerges when comparing GT and low-testing students, with 9.11% in the bottom quintile for both groups and 10.34% in the top. In ELA, between 8% and 10% are similarly classified in the highest or lowest quintile across groups.

While many teachers fall into the same or adjacent VAM quintiles across student groups, a meaningful portion exhibit substantial differences in effectiveness. For instance, 10.84% of math teachers and 14.25% of ELA teachers have a high-testing VAM that is at least two quintiles higher than their low-testing VAM.[4] When comparing GT and low-testing VAMs, these shares are even larger — 13.13% in math and 15.93% in ELA — indicating that more teachers exhibit a relative advantage in teaching GT students than high-testing students, despite the substantial overlap between the groups. This pattern reinforces the value of examining both GT and high-achieving students separately when assessing teacher effectiveness.

The bottom left and top right bins in each panel indicate major relative advantages in VAMs across group. Teachers who exhibit high levels of effectiveness for one group while being ineffective for the other are relatively rare. Only 0.61% of math teachers and 1.16% of ELA teachers are highly effective for GT students but ineffective for low-testing students, while 0.55% of math teachers and 1.932% of ELA teachers fall into the category of being highly effective for low-testing students but ineffective for GT students. A similar pattern emerges when comparing high- and low-testing students. These findings reinforce the presence of significant variability in teacher effectiveness

[4] In an analogous version of Table 4 constructed using the randomized VAMs, the share of teachers classified as having a substantial difference in effectiveness is significantly smaller at 5.37% in math and 9.27% in ELA. This further supports the interpretation that the observed variation reflects meaningful differences in subgroup-specific effectiveness rather than random error.

across different student subgroups.

5 RQ2: Are observable teacher characteristics associated with higher VAMs for GT, high-, or low-testing students, and if so, do these characteristics differ between the groups?

Although the current literature suggests most observable teacher characteristics are not good predictors of teacher effectiveness as measured by VAMs (Kane and Staiger, 2008; Chingos and Peterson, 2011; Loeb et al., 2014; Master et al., 2016; Wood et al., 2024), this may be attributable to VAMs being calculated across all students, instead of subgroups. We therefore examine whether observable teacher characteristics are more predictive of effectiveness when VAMs are calculated separately for GT, high-testing, and low-testing students.

To do this, we use the following standard linear regression model to estimate the association between observable teacher characteristics and VAMs by subject and group:

$$VAM_{jt}^{subject,group} = \beta_0 + X_{jt}B_x + Z_{jt}B_z + \theta_s + \theta_t + \epsilon_{jt}, \quad (3)$$

where VAM_{jt} is a standardized VAM calculated from Equation (1) for each subject and group; X_{jt} is a vector of teacher characteristics including having a Master's or PhD, National Board Certification, years of experience, gender, and race; Z_{jt} is a vector of classroom characteristics and controls for the share of non-White/non-Asian, EL, FRL, and GT students within each classroom; and θ_s and θ_t are school and year fixed effects, respectively, that account for potential different shocks across schools and years. Using Equation (3), we estimate a separate regression for each subgroup and subject to describe whether the observable characteristics of effective teachers differ across student groups.

Table 5 contains the regression results from Equation (3) with each column displaying the association between teacher characteristics and subgroup-specific VAMs by subject. Overall, few observable characteristics are strongly predictive of teacher effectiveness, consistent with prior research. However, a few patterns are notable.

First, teacher experience is largely uncorrelated with effectiveness, with the exception of late-career teachers. Teachers with at least 10 years of experience have significantly lower math VAMs for both high-testing (–0.180 SD) and GT students (–0.208 SD) compared to novice teachers. The estimate for low-testing students is similarly negative (–0.137 SD) but not statistically significant. This contrasts with prior findings that teacher effectiveness tends to rise with experience (Podolsky et al., 2019), though that relationship may not hold for subgroup-specific VAMs (Wood et al., 2024).

The negative association in later years may reflect diminishing returns to experience, changes in pedagogical alignment, or reduced adaptability to evolving instructional strategies and student needs. For GT students specifically, it is also possible that more experienced teachers received less training in differentiated instruction for gifted learners. The National Association for Gifted Children (NAGC) did not publish its first formal guidelines until 1998, and targeted investments in GT training became more widespread only with the expansion of the Jacob Javits Gifted and Talented Students Education Act under No Child Left Behind (National Association for Gifted Children, 2005). Many of the late-career teachers in our sample had already entered the profession before these supports were widely available, leaving them less prepared and confident in meeting the needs of GT students — a pattern that prior research suggests may contribute to their lower effectiveness for GT students (Berman et al., 2012; Klassen and Tze, 2014; Matheis et al., 2017). However, given the similar estimates for high- and low-testing students, broader explanations — such as reduced instructional flexibility in later stages of teaching — may also contribute to lower effectiveness across all subgroups.

In contrast, some teacher characteristics are positively associated with subgroup-specific VAMs. Female teachers and those with National Board Certification are associated with higher ELA VAMs for low-testing students (0.101 and 0.141 SDs, respectively). We also find that Latino/a and Asian teachers tend to be more effective than White teachers, though the coefficients for the GT-specific VAMs are not significant. Given the large share of Latino/a students in LAUSD, this may reflect a positive student–teacher race match effect (Dee, 2004; Egalite et al., 2015; Joshi et al., 2018; Wood and Lai, 2022), but further analysis of this hypothesis is beyond the scope of this work.

To complement these results, we estimate a similar model using DVAMs as the outcome. These results, reported in Appendix Table A.1, suggest that observable teacher characteristics are not strong predictors of relative effectiveness across student subgroups. The only statistically significant association is between teacher gender and GT DVAM in math: female teachers are estimated to be 2.41 percentile points more effective, on average, with GT students relative to low-testing students.

6 RQ3: Are GT and high-testing students assigned to classes taught by teachers with higher GT or high-testing VAMs?

To explore how students are assigned to teachers, we regress each student's teacher's VAM on a classification indicator (GT, non-GT high-testing, and each of the GT categories) using the following equation:

$$Y_{it}^{subject,group} = \alpha_0 + \alpha_1 Classification_{it} + \theta_{sgt} + \epsilon_{it}, \tag{4}$$

where $Y_{it}^{subject,group}$ represents a teacher's VAM for a given subject and group of students in year t; $Classification$ is an indicator equal to one if student i is in the specified category in year t; θ_{sgt} is a school-grade-year fixed effect that allows us to focus on student distribution within individual school-grade-year combinations; and α_1 is the coefficient of interest interpreted as the expected difference in teacher VAMs for students in each category compared to students not in that category. We also conduct analysis separately for elementary (grades 3-5) and middle school (grades 6-8) to determine if assigning GT students to more effective teachers is emphasized more in later grades where tracking is more common.

Panel A of Table 6 contains results from Equation (4), showing that GT students are consistently assigned to teachers with higher VAMs across all student groups. In math, GT students are matched to teachers whose GT-specific VAMs are 0.067 SDs higher, high-testing VAMs are 0.058 SDs higher, and low-testing VAMs are 0.028 SDs higher, on average. A similar pattern holds for ELA, with assignment differences of 0.054, 0.047, and 0.018 SDs, respectively. Importantly, the difference between the relevant subgroup VAMs (GT and high-testing) and the low-testing VAMs is statistically significant in both subjects, suggesting that GT students are not just placed with generally effective teachers, but with teachers who appear particularly effective at instructing students like them. This pattern provides suggestive evidence of targeted matching between students and teachers based on subgroup-specific instructional effectiveness.

Non-GT high-testing students are also assigned to teachers with higher VAMs, with effects ranging from 0.017 SDs for low-testing ELA VAMs to 0.066 SDs for GT math VAMs. Within the GT subgroups, we observe consistent and substantial positive associations with teacher VAMs across both subjects. HG students, for example, are assigned to teachers whose GT VAMs are 0.084 SDs higher in math and 0.186 SDs higher in ELA — the largest subgroup differences observed. With the exception of students gifted in the arts, students in all other GT classifications are matched with teachers who have significantly higher GT VAMs, ranging from 0.045 to 0.066 SDs in math and 0.041 to 0.046 SDs in ELA. In all cases, the GT VAM is statistically higher than the corresponding low-testing VAM, reinforcing that subgroup-specific assignment patterns extend to GT subgroups as well.

Panel B Table 6 presents a breakdown of the GT estimates from Panel A by school level. In middle school (grades 6–8), GT students are consistently assigned to teachers with significantly higher VAMs across all subject–group combinations. For example, GT students are matched with teachers whose math VAMs are 0.092 SDs higher for GT students and 0.081 SDs higher for high-testing students. In ELA, the corresponding differences are 0.065 and 0.054 SDs, respectively. These VAMs are again significantly larger than the corresponding low-testing VAMs.

In contrast, the pattern in elementary school (grades 3–5) is less consistent. While GT students are assigned to teachers with significantly higher ELA VAMs — 0.021 SDs for GT VAMs and 0.022 SDs for GT and high-testing VAMs — there are no significant differences in math teacher assignment across any of the three VAM measures.

This disparity likely reflects improved tracking capabilities in middle schools, which enable better matching of GT students with effective teachers. These findings suggest that targeted assignment to teachers with subgroup-specific strengths becomes more pronounced in later grades, particularly in math. Such matching may play a critical role in supporting students' academic development and highlights the importance of intentional teacher assignments as students progress through school.

Additionally, we explore possible heterogeneity by FRL eligibility, EL status, non-White/non-Asian status, school FRL percentage (90%+ FRL enrollment vs. <50% eligible),[5] and school enrollment size (top quartile vs. bottom quartile) by interacting the GT indicator with these characteristics (represented by X in the equation below):

$$Y_{it}^{subject,group} = \alpha_0 + \alpha_1 GT_{it} X_{it} + \alpha_2 GT_{it} + \alpha_3 X_{it} + \theta_{sgt} + \epsilon_{it} \tag{5}$$

Table 7 presents a summary of results from Equation (5) in which we examine heterogeneity across the aforementioned student and school characteristics.[6] Each panel is structured as follows: the first row provides an estimate of the difference in teacher VAMs between GT and non-GT students who share the given characteristic. The second row offers a similar comparison

[5] We chose 50% and 90% as cut-offs to compare schools with significantly different student populations. Since more than half of all schools in LAUSD have 78-86% of students eligible for FRL, these thresholds provide a more clear distinction between schools than comparing schools in the top and bottom quartiles of FRL enrollment.

[6] See Table A.2 in the Appendix for full results.

for GT and non-GT students without the characteristic of interest. The third row shows the difference between the previous two rows, highlighting differential teacher assignment for GT students based on whether they possess the given characteristic.

Panel A shows that GT students are consistently assigned to teachers with higher VAMs regardless of FRL eligibility. FRL-eligible GT students are assigned teachers whose math GT VAMs are 0.068 SDs higher than those of comparable non-GT peers, while non-FRL GT students receive teachers with math GT VAMs 0.061 SDs higher than non-GT students ineligible for FRL. The difference (0.007 SDs) is not statistically significant. However, for low-testing VAMs in ELA, FRL-eligible GT students receive teachers with VAMs 0.024 SDs higher than their non-GT peers, compared to 0.003 SDs among non-FRL students — a statistically significant 0.022 SD difference. This pattern suggests that FRL-eligible GT students are more likely to be matched with teachers who have higher VAMs for low-testing students, even though these GT students are typically high-testers.

Panel B reveals a different pattern by EL status. While non-EL GT students are consistently matched with teachers who have higher VAMs than their non-GT peers, EL GT students are not. In fact, EL GT students are assigned to teachers with math GT VAMs that are 0.043 SDs lower and ELA GT VAMs that are 0.060 SDs lower than those of non-EL GT students. These differences suggest that EL GT students are less likely to be matched with teachers who are particularly effective for GT students. Despite their academic qualifications, EL GT students may therefore face additional barriers to accessing high-quality instruction.

Panel C compares assignment patterns by student race/ethnicity. Differences in access to effective teachers by GT status are broadly similar across groups, with no significant differences across any subgroup VAMs or subject. Panel D disaggregates GT teacher assignment by school-level FRL concentration. GT students in schools with over 90% FRL enrollment are assigned teachers with significantly higher VAMs in both math (0.053 to 0.077 SDs) and ELA (0.051 to 0.059 SDs) compared to non-GT students in the same schools. In contrast, GT students in schools with less than 50% FRL enrollment are not consistently matched to significantly more effective teachers. Notably, the differences in GT assignment between high- and low-FRL schools are statistically significant for low-testing math and ELA VAMs: GT students in high-FRL schools are placed with teachers whose low-testing VAMs are 0.057 and 0.095 SDs higher, respectively, than those assigned to GT students in low-FRL schools.

Panel E shows that in the smallest schools (bottom enrollment quartile), GT students are not assigned to teachers with significantly different VAMs across any subject. In contrast, GT students in the largest schools (top enrollment quartile) are matched with teachers who are substantially more effective in math: compared to non-GT students in the same schools, GT students have teachers with GT VAMs 0.128 SDs higher and high-testing VAMs 0.122 SDs higher. The differences in ELA VAMs are smaller (0.004 to 0.077 SDs). As a result, GT students in smaller schools are assigned to teachers with GT math VAMs 0.133 SDs lower than GT students in larger schools, highlighting meaningful disparities in access to effective math instruction based on school size.

7 RQ4: Are GT students assigned to teachers who are relatively more effective at teaching GT students?

We next examine whether GT students are matched with teachers who are relatively more effective in teaching GT students than low-testing students. To assess this, we use the DVAM measure that compares GT and low-testing VAMs and construct several relative advantage indicators: whether a teacher has any relative advantage (> 0), at least a 10 percentile point advantage (> 10), or a 20 percentile point advantage (> 20) in teaching GT students. These indicators allow us to examine the extent to which teacher assignment aligns with subgroup-specific relative advantage, regardless of the teacher's absolute VAM level.

Table 8 presents the results from this analysis. Among all GT students, we find consistent and statistically significant evidence that they are more likely to be assigned to teachers with a relative advantage in teaching GT students. In math, GT students are 1.7 percentage points (pp) more likely to be matched with teachers with a relative advantage, 1.6 pp more likely to be assigned to teachers with at least a 10 percentile point advantage, and 0.9 pp more likely to have a teacher with at least a 20 percentile point advantage in teaching GT students. In ELA, the differences are generally smaller (1.0-1.3 pp), but remain statistically significant across all thresholds. These results suggest that GT students are systematically assigned to math teachers with a demonstrated comparative advantage in teaching students like them.

Non-GT high-testing students are also more likely to be assigned to teachers with a relative advantage in teaching GT students, with differences of 0.6 to 1.1 pp in math and up to 1.1 pp in ELA. These patterns are comparable in magnitude to those observed for GT students, suggesting that assignment to teachers with subgroup-specific strengths is not limited to formally identified GT students. However, this broader pattern of assignment masks meaningful variation across GT subgroups. Students classified as GT through intellectual ability (IA), high achievement (HA), or math-specific achievement (Math) are significantly more likely to be placed with teachers who have a relative advantage in teaching GT students across both subjects, with differences are as large as 1.8 pp in math and 1.0 pp in ELA. In contrast, students identified as gifted in ELA or the arts do not appear to be matched with relatively more effective teachers. The most pronounced effects are observed for GT–HG students, who are 5.7 pp more likely to be assigned an ELA teacher with a relative advantage in teaching GT students, and 4.2

pp more likely to be assigned an ELA teacher whose GT advantage exceeds 20 percentile points. Appendix Table A.3 replicates this analysis using relative advantage in teaching high-testing (compared to low-testing) students. Results are nearly identical, reinforcing the conclusion that GT and high-achieving students are similarly assigned to teachers with subgroup-specific instructional strengths.

Taken together, these findings suggest that there is an effort to place GT students with teachers who have a relative advantage in teaching them. Although the magnitudes are modest — generally between 1 and 2 pp — they are consistent across multiple GT subgroups and statistically significant for several. Combined with the earlier results in Table 6, which show that GT students are assigned to teachers with higher overall VAMs, this evidence indicates that GT students are not just placed with generally effective teachers, but to some extent, with teachers who are relatively more effective at teaching GT students. This subtle but meaningful differentiation in teacher assignment highlights the potential for using subgroup-specific instructional effectiveness to better match students with the educators best equipped to support their learning.

8 Robustness Checks

To assess the sensitivity of our main results (Table 6) we conduct a series of robustness checks using alternative sample restrictions and subgroup definitions. These results are reported in Appendix Tables A.4 through A.7.

Appendix Table A.4 re-estimates the models in Table 6 using a relaxed inclusion threshold of 10 students per teacher–subgroup pair (instead of the baseline 15-student threshold). Appendix Table A.5 increases this threshold to 20 students. Appendix Table A.6 retains the 15-student threshold, but modifies the high- and low-testing definitions to include only students in the top and bottom terciles (rather than quartiles) of prior-year test scores. Appendix Table A.7 changes the definition the high- and low-testing definitions to students in the top and bottom quintiles.

Across all specifications, we find that the core results hold: GT students are consistently assigned to teachers with higher VAMs, and these assignments are significantly larger for GT- and high-testing-specific VAMs than for low-testing VAMs. Effect sizes for GT math and ELA VAMs under these alternative thresholds and definitions range from 0.065 to 0.068 SDs and 0.052 to 0.056 SDs, respectively — nearly identical to the corresponding estimates in Table 6 (0.067 and 0.054 SDs). These consistent results across multiple specifications reinforce the robustness of our main findings.

9 Discussion and Conclusion

This paper utilizes data from LAUSD to investigate whether certain teachers are more effective in fostering academic growth among GT and high-testing students compared to their lower-testing peers. It also examines whether observable teacher characteristics predict subgroup-specific effectiveness and evaluates the extent to which GT students are systematically placed with teachers who are particularly effective at teaching them.

Our analysis reveals that a subset of teachers displays a sizeable relative advantage in instructing GT and high-testing students. However, no observable teacher characteristics are consistently associated with higher effectiveness for GT students. We find that GT students are assigned to teachers with higher GT-specific, high-testing, and low-testing VAMs, but the assignment effect is significantly larger for the GT-specific VAM than for the low-testing VAM. This pattern holds across nearly all GT subgroups and is also present for non-GT high-testing students. Moreover, while most teacher assignments appear to reflect general effectiveness, there is modest evidence that GT students are more likely to be assigned to teachers with a relative advantage in teaching them. Assignment disparities by school size and EL status further suggest that access to subgroup-aligned instruction is not uniformly distributed. By estimating separate VAMs for GT, high-testing, and low-testing students, this paper contributes to the literature on subgroup-specific teacher effectiveness and, to our knowledge, is the first to construct a GT-specific VAM to assess both teacher effectiveness and assignment alignment for gifted students.

These findings carry several policy implications. First, although GT students are generally assigned to teachers with higher VAMs, current assignment patterns may not fully capitalize on opportunities to match students with teachers who have demonstrated relative strengths in teaching them. The small magnitude of our estimates suggests that while subgroup-specific alignment exists, it may not be a central consideration in assignment decisions. Additionally, we find that EL classification influences teacher placement for GT students, raising concerns about equitable access to tailored instruction. Together, these patterns suggest the potential for more deliberate matching, though evaluating the feasibility and impact of such efforts lies beyond the scope of this study.

Second, the absence of strong associations between observable teacher characteristics and effectiveness with GT students highlights the difficulty of identifying teachers who are particularly well-suited to support this group. While this may reflect the role of unobservable attributes, it also underscores the challenge administrators face in allocating effective teachers to GT students based on readily available information. One strategy to address this challenge is the development and use of subgroup-specific value-added or growth measures that more directly capture teachers' demonstrated effectiveness with these students. These metrics could inform staffing decisions and support more deliberate student–teacher matching. In parallel, school districts may consider investing in specialized professional development to equip a broader pool of teachers with evidence-based strategies for differentiating instruction and meeting the academic and social-emotional needs of

GT learners.

Finally, the finding that GT students in small schools do not benefit from higher math VAM teachers as much as their peers in larger schools highlights a potential inequity in access to effective instruction. Policymakers should consider how resources are distributed across schools of different sizes and demographics. Ensuring that small and potentially under-resourced schools have access to highly-effective teachers, particularly in subjects like mathematics, is crucial for maintaining equity in GT education. This could involve targeted recruitment efforts, incentives for highly-effective teachers to work in smaller or under-resourced schools, or district-wide initiatives that share effective teaching practices across schools.

Table 1: GT and High-Testing, Non-GT Students Descriptive Characteristics

Panel A: All Students	Top Quartile	Female	FRL	EL	SWD	Asian	Black	Latino/a	White	Other Race	Math Test	ELA Test
Non-GT, Non-High-Testing	0.000	0.490	0.852	0.296	0.131	0.022	0.089	0.808	0.061	0.021	-0.332	-0.373
N = 1,750,319	(0.000)	(0.500)	(0.355)	(0.457)	(0.337)	(0.145)	(0.285)	(0.394)	(0.239)	(0.143)	(0.793)	(0.796)
GT	0.700	0.490	0.643	0.010	0.013	0.109	0.055	0.581	0.202	0.053	1.215	1.108
N = 422,592	(0.458)	(0.500)	(0.479)	(0.099)	(0.112)	(0.312)	(0.229)	(0.493)	(0.402)	(0.224)	(0.869)	(0.783)
High-Testing, Non-GT	1.000	0.530	0.721	0.034	0.033	0.061	0.068	0.684	0.143	0.044	0.731	0.731
N = 318,445	(0.000)	(0.499)	(0.448)	(0.182)	(0.179)	(0.240)	(0.251)	(0.465)	(0.350)	(0.206)	(0.739)	(0.648)
Panel B: Analytical Sample	Top Quartile	Female	FRL	EL	SWD	Asian	Black	Latino/a	White	Other Race	Math Test	ELA Test
GT	0.786	0.488	0.689	0.006	0.011	0.084	0.056	0.636	0.171	0.053	1.175	1.032
N = 214,085	(0.410)	(0.500)	(0.463)	(0.079)	(0.106)	(0.277)	(0.230)	(0.481)	(0.377)	(0.225)	(0.883)	(0.748)
High-Testing, Non-GT	1.000	0.542	0.728	0.015	0.030	0.058	0.068	0.686	0.141	0.048	0.749	0.747
N = 153,669	(0.000)	(0.498)	(0.445)	(0.122)	(0.172)	(0.233)	(0.251)	(0.464)	(0.348)	(0.213)	(0.730)	(0.607)

Table 2: GT Classifications

Classification	Criteria	Evaluation Method	Eligible Grades
Intellectual Ability (IA)	99.5-99.4 percentile on diagnostic exam.	Test administered by LAUSD GATE psychologist.	2nd semester of kindergarten and above.
Highly-Gifted (HG)	99.5+ percentile on diagnostic exam.	Test administered by LAUSD GATE psychologist.	2nd semester of kindergarten and above.
High Achievement (HA)	High achievement levels in ELA and math for two consecutive years.	Smarter Balanced Assessment, district-approved standardized tests, and/or class grades.	5th grade and above.
Math	High achievement levels in math for three consecutive years.	Smarter Balanced Assessment, district-approved standardized tests, and/or class grades.	5th grade and above.
Language Arts	High achievement levels in ELA for three consecutive years.	Smarter Balanced Assessment, district-approved standardized tests, and/or class grades.	5th grade and above.
Arts (Performing and Visual)	High levels in dance, voice, drama, drawing, or painting.	Student audition, portfolio, or demonstration.	2nd grade and above.
Creative Ability	High levels of problem-solving, originality, and/or imagination.	Student portfolio.	2nd grade and above.
Leadership Ability	Passion about community/ environmental issue, exceptional problem-solving, and sense of purpose.	Student portfolio.	2nd grade and above.

Figure 1: Distribution of GT Students Across and Within Schools

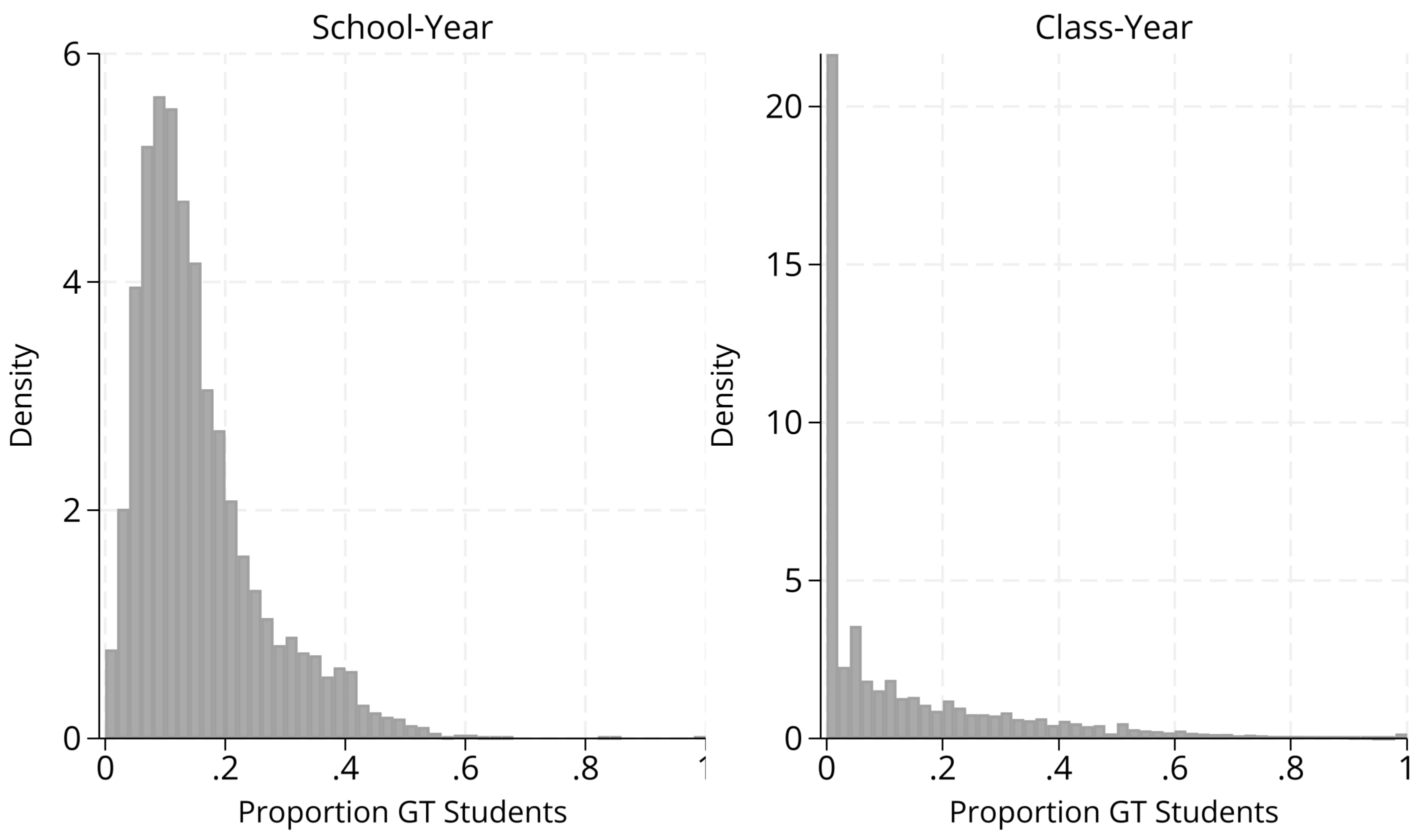

Table 3: Student and Teacher Descriptive Characteristics

Panel A: Student	Share of Students	Female	FRL	EL	SWD	Asian	Black	Latino/a	White	Other Race	Math Test	ELA Test
<u>**Student-Year**</u>												
All Students	1.000	0.506	0.781	0.108	0.063	0.043	0.079	0.738	0.105	0.036	0.265	0.208
N = 917,218		(0.500)	(0.413)	(0.311)	(0.242)	(0.202)	(0.270)	(0.440)	(0.306)	(0.185)	(0.993)	(0.924)
Non-GT	0.767	0.511	0.810	0.139	0.078	0.030	0.086	0.769	0.084	0.030	-0.009	-0.044
N = 703,191		(0.500)	(0.393)	(0.346)	(0.269)	(0.171)	(0.281)	(0.422)	(0.278)	(0.171)	(0.851)	(0.821)
GT	0.233	0.488	0.689	0.006	0.012	0.083	0.056	0.636	0.171	0.053	1.175	1.032
N = 214,027		(0.500)	(0.463)	(0.078)	(0.107)	(0.276)	(0.230)	(0.481)	(0.376)	(0.224)	(0.882)	(0.747)
GT - Intellectual Ability	0.100	0.482	0.693	0.009	0.013	0.088	0.057	0.630	0.171	0.055	1.091	0.986
N = 92,137		(0.500)	(0.461)	(0.094)	(0.113)	(0.283)	(0.232)	(0.483)	(0.376)	(0.228)	(0.909)	(0.766)
GT - Math	0.062	0.423	0.778	0.006	0.012	0.066	0.041	0.742	0.105	0.046	1.112	0.767
N = 56,487		(0.494)	(0.416)	(0.074)	(0.109)	(0.248)	(0.199)	(0.438)	(0.307)	(0.209)	(0.821)	(0.652)
GT - High Achievement	0.058	0.528	0.620	0.002	0.007	0.096	0.061	0.571	0.214	0.058	1.395	1.330
N = 53,249		(0.499)	(0.485)	(0.049)	(0.085)	(0.294)	(0.239)	(0.495)	(0.410)	(0.234)	(0.845)	(0.685)
GT - Language Arts	0.008	0.695	0.589	0.002	0.008	0.059	0.076	0.557	0.252	0.056	1.063	1.377
N = 7,583		(0.461)	(0.492)	(0.040)	(0.091)	(0.236)	(0.266)	(0.497)	(0.434)	(0.229)	(0.770)	(0.574)
GT - Arts	0.003	0.739	0.480	0.011	0.039	0.053	0.196	0.351	0.373	0.028	0.783	0.885
N = 2,653		(0.439)	(0.500)	(0.102)	(0.193)	(0.224)	(0.397)	(0.477)	(0.484)	(0.164)	(1.001)	(0.848)
GT - Highly-Gifted	0.002	0.407	0.484	0.007	0.019	0.200	0.051	0.360	0.322	0.066	1.906	1.634
N = 1,870		(0.491)	(0.500)	(0.083)	(0.137)	(0.400)	(0.220)	(0.480)	(0.468)	(0.249)	(1.059)	(0.858)

Panel B: Teacher			Experience									
	Master's or PhD	National Board Certification	Novice (1-2 yrs)	Early (3-5 yrs)	Middle (6-10 yrs)	Late (10+ yrs)	Female	Asian	Black	Latino/a	White	Other Race
<u>**Student-Year**</u>												
Non-GT	0.389	0.049	0.023	0.077	0.168	0.732	0.661	0.101	0.120	0.304	0.436	0.040
N = 702,733	(0.487)	(0.216)	(0.149)	(0.266)	(0.374)	(0.443)	(0.473)	(0.301)	(0.325)	(0.460)	(0.496)	(0.195)
GT	0.398	0.066	0.014	0.061	0.165	0.760	0.691	0.110	0.105	0.258	0.484	0.043
N = 214,171	(0.489)	(0.248)	(0.119)	(0.239)	(0.371)	(0.427)	(0.462)	(0.313)	(0.306)	(0.438)	(0.500)	(0.202)

Sample contains students paired with teachers with at least 15 students in the top and bottom quartile of the previous year's test scores and 15 GT students. GT (gifted and talented student). High-testing (student in top quartile of the previous year's math or language arts test score distribution). FRL (free or reduced-price lunch). EL (English learner). SWD (student with disability).

Figure 2: High-Testing vs. Low-Testing and GT vs. Low-Testing VAM Scatterplots

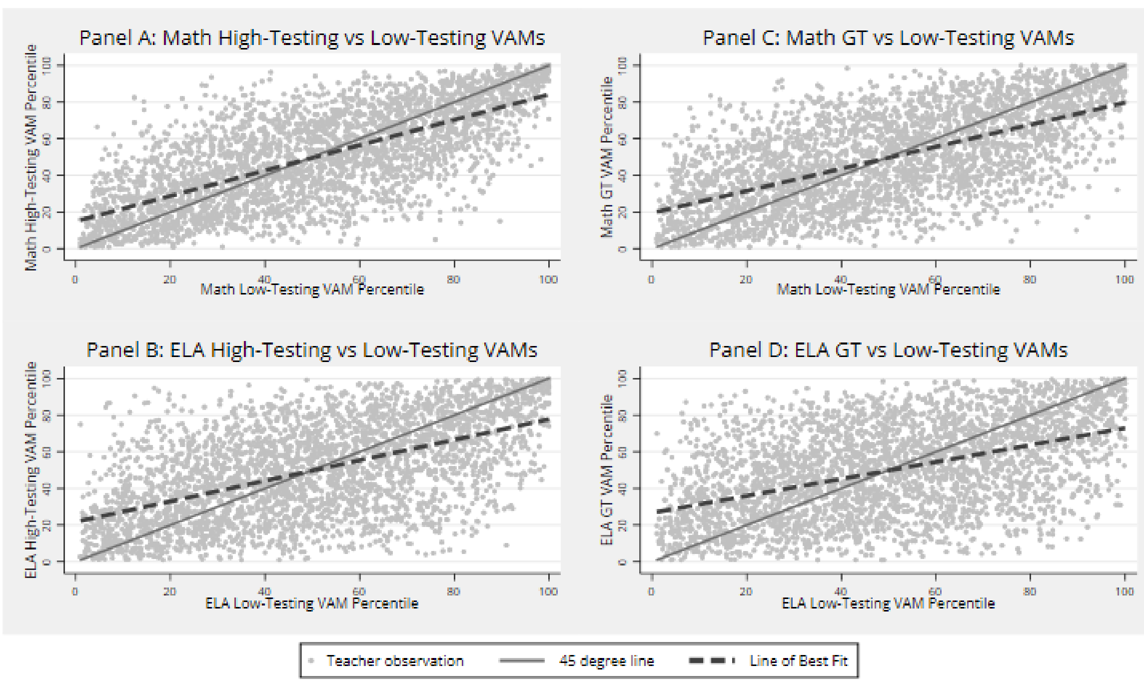

Figure 3: DVAM and Randomized DVAM Distributions

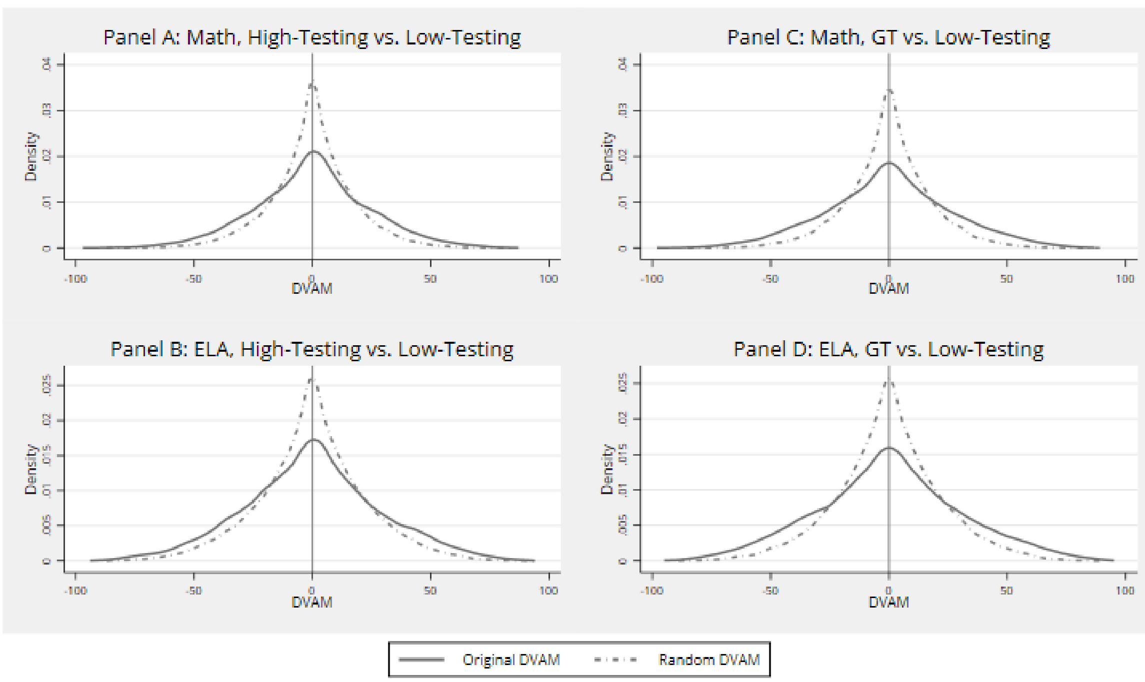

Table 4: Distribution of High-Testing, GT, and Low-Testing VAMs by Quintile

High-Testing and Low-Testing VAM Quintiles						
Panel A: Math		**Low-Testing VAM Quintile**				
		1	**2**	**3**	**4**	**5**
High-Testing VAM Quintile	1	**10.30%**	5.25%	2.89%	1.24%	0.36%
	2	5.42%	**5.62%**	4.81%	3.02%	1.13%
	3	2.74%	4.92%	**5.28%**	4.67%	2.41%
	4	1.19%	3.19%	4.71%	**6.09%**	4.82%
	5	0.39%	1.01%	2.32%	4.99%	**11.28%**
Panel B: ELA		**Low-Testing VAM Quintile**				
		1	**2**	**3**	**4**	**5**
High-Testing VAM Quintile	1	**8.89%**	4.64%	3.21%	2.34%	0.95%
	2	5.15%	**5.22%**	4.41%	3.42%	1.80%
	3	3.23%	4.82%	**4.78%**	4.41%	2.78%
	4	1.91%	3.67%	4.67%	**5.09%**	4.65%
	5	0.85%	1.64%	2.95%	4.74%	**9.81%**
GT and Low-Testing VAM Quintiles						
Panel C: Math		**Low-Testing VAM Quintile**				
		1	**2**	**3**	**4**	**5**
GT VAM Quintile	1	**9.11%**	5.20%	3.48%	1.70%	0.55%
	2	5.44%	**5.39%**	4.33%	3.21%	1.62%
	3	3.16%	4.54%	**5.02%**	4.58%	2.72%
	4	1.71%	3.54%	4.41%	**5.58%**	4.75%
	5	0.61%	1.33%	2.78%	4.92%	**10.34%**
Panel D: ELA		**Low-Testing VAM Quintile**				
		1	**2**	**3**	**4**	**5**
GT VAM Quintile	1	**8.05%**	4.68%	3.42%	2.55%	1.32%
	2	4.95%	**5.08%**	4.39%	3.46%	2.12%
	3	3.50%	4.48%	**4.65%**	4.50%	2.88%
	4	2.36%	3.73%	4.40%	**4.83%**	4.67%
	5	1.16%	2.03%	3.15%	4.67%	**8.99%**

Each teacher-year observation is assigned a quintile by its relative rank in the VAM distribution. The lowest quintile is “1”, while the highest is “5”.

Table 5: Association between Teacher Characteristics and VAMs by Student Group

	Math VAM			ELA VAM		
	(1) High-Testing	(2) GT	(3) Low-Testing	(4) High-Testing	(5) GT	(6) Low–Testing
National Board Certification	0.063	0.030	0.115	0.075	0.055	0.141*
	(0.071)	(0.074)	(0.077)	(0.069)	(0.070)	(0.070)
Master's	-0.019	0.006	-0.021	0.063	0.037	0.052
	(0.032)	(0.033)	(0.031)	(0.033)	(0.033)	(0.033)
Exp: Early Career (3-5 yrs)	0.070	0.023	0.088	0.080	0.018	0.010
	(0.062)	(0.066)	(0.065)	(0.066)	(0.067)	(0.063)
Exp: Middle Career (6-9 yrs)	-0.005	-0.030	0.007	0.036	-0.035	-0.095
	(0.068)	(0.071)	(0.068)	(0.066)	(0.070)	(0.064)
Exp: Late Career (10+ yrs)	-0.180*	-0.208**	-0.137	-0.023	-0.081	-0.121
	(0.071)	(0.074)	(0.072)	(0.071)	(0.078)	(0.063)
Female	-0.005	-0.033	0.027	0.077*	0.057	0.101**
	(0.037)	(0.039)	(0.037)	(0.036)	(0.035)	(0.038)
Asian	0.141*	0.125	0.134*	0.091	0.048	0.026
	(0.058)	(0.064)	(0.061)	(0.062)	(0.065)	(0.064)
Black	-0.117	-0.082	-0.128	-0.096	-0.101	-0.058
	(0.075)	(0.079)	(0.069)	(0.061)	(0.069)	(0.063)
Latino/a	0.075	0.062	0.134**	0.074	0.059	0.101*
	(0.048)	(0.048)	(0.047)	(0.043)	(0.043)	(0.048)
Other Race	-0.020	-0.005	-0.060	-0.019	-0.007	-0.045
	(0.083)	(0.087)	(0.076)	(0.095)	(0.101)	(0.087)
N	16,784	16,784	16,784	16,504	16,504	16,504

School-clustered standard errors are shown in parentheses. All columns include school and year fixed effects and classroom characteristics. Reference groups are novice teachers (1-2 years) and white teachers. *p<.05. **p<.01. ***p<.001.

Table 6: Association Between GT Status & VAMs

	Panel A: Overall					
	Math VAM			ELA VAM		
	(1) High-Testing	(2) GT	(3) Low-Testing	(4) High-Testing	(5) GT	(6) Low–Testing
GT	0.058***	0.067***	0.028***	0.047***	0.054***	0.018*
	(0.010)†	(0.012)†	(0.008)	(0.009)†	(0.010)†	(0.009)
Non-GT High-Testing	0.060***	0.066***	0.040***	0.043***	0.045***	0.017*
	(0.009)†	(0.010)†	(0.008)	(0.006)†	(0.007)†	(0.007)
GT - IA	0.039***	0.045***	0.021**	0.036***	0.041***	0.015
	(0.008)†	(0.009)†	(0.007)	(0.008)†	(0.008)†	(0.008)
GT - Math	0.050***	0.058***	0.021*	0.034***	0.041***	0.012
	(0.010)†	(0.011)†	(0.008)	(0.008)†	(0.008)†	(0.007)
GT - HA	0.058***	0.066***	0.026***	0.038***	0.046***	0.017
	(0.010)†	(0.012)†	(0.008)	(0.009)†	(0.010)†	(0.009)
GT - LA	0.046***	0.052***	0.039***	0.044***	0.046***	0.015
	(0.011)†	(0.012)†	(0.009)	(0.010)†	(0.011)†	(0.012)
GT - Arts	0.010	0.006	-0.009	0.010	-0.012	0.012
	(0.017)†	(0.019)	(0.013)	(0.019)	(0.026)†	(0.018)
GT - HG	0.071*	0.084*	0.021	0.152***	0.186***	0.000
	(0.030)†	(0.037)†	(0.027)	(0.036)†	(0.043)†	(0.030)
N	880,849	880,849	880,849	902,808	902,808	902,808

	Panel B: By School Level					
	Math VAM			ELA VAM		
	(1) High-Testing	(2) GT	(3) Low-Testing	(4) High-Testing	(5) GT	(6) Low–Testing
Elementary (3-5)	0.002	-0.001	0.001	0.022***	0.021***	0.017***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Middle (6-8)	0.079***	0.092***	0.038***	0.054***	0.065***	0.019
	(0.014)	(0.016)	(0.011)	(0.012)	(0.013)	(0.012)
N	880,849	880,849	880,849	902,808	902,808	902,808

Teacher-clustered standard errors shown in parentheses. All columns include school-grade-year-fixed effects. *p<.05. **p<.01. ***p<.001. † denotes difference in estimated association is statistically different (p<.001) from the low-testing estimate.

Table 7: Association Between GT Status & VAMs, Heterogeneity

	Math			ELA		
	High-Testing	GT	Low-Testing	High-Testing	GT	Low-Testing
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: FRL						
FRL	0.059***	0.068***	0.034***	0.048***	0.053***	0.024**
	(0.011)	(0.012)	(0.009)	(0.009)	(0.009)	(0.009)
Non-FRL	0.054***	0.061***	0.010	0.044***	0.056***	0.003
	(0.013)	(0.016)	(0.009)	(0.013)	(0.015)	(0.013)
Difference	0.004	0.007	0.024*	0.003	-0.003	0.022*
	(0.011)	(0.014)	(0.009)	(0.011)	(0.013)	(0.011)
N	880,849	880,849	880,849	902,808	902,808	902,808
Panel B: EL						
EL	0.028	0.020	0.019	-0.000	-0.007	0.028
	(0.017)	(0.019)	(0.016)	(0.016)	(0.017)	(0.017)
Non-EL	0.055***	0.063***	0.025**	0.046***	0.053***	0.016
	(0.010)	(0.011)	(0.008)	(0.009)	(0.010)	(0.009)
Difference	-0.028	-0.043*	-0.006	-0.046**	-0.060***	0.012
	(0.017)	(0.018)	(0.016)	(0.017)	(0.018)	(0.018)
N	880,849	880,849	880,849	902,808	902,808	902,808
Panel C: Non-White, Non-Asian						
Non-White, Non-Asian	0.057***	0.064***	0.030***	0.047***	0.052***	0.022*
	(0.010)	(0.011)	(0.008)	(0.009)	(0.009)	(0.009)
White or Asian	0.057***	0.068***	0.016	0.038**	0.057***	-0.000
	(0.015)	(0.018)	(0.011)	(0.013)	(0.016)	(0.014)
Difference	-0.000	-0.004	0.013	0.009	-0.005	0.023
	(0.013)	(0.016)	(0.010)	(0.012)	(0.014)	(0.012)
N	880,849	880,849	880,849	902,808	902,808	902,808
Panel D: 90%+ FRL School						
90%+ FRL School	0.062**	0.077***	0.053**	0.059**	0.057**	0.051**
	(0.020)	(0.022)	(0.019)	(0.018)	(0.020)	(0.019)
<50% FRL School	0.068*	0.055	-0.004	0.014	0.030	-0.044
	(0.031)	(0.035)	(0.022)	(0.028)	(0.031)	(0.030)
Difference	-0.006	0.022	0.057*	0.045	0.027	0.095**
	(0.037)	(0.042)	(0.029)	(0.033)	(0.037)	(0.035)
N	486,117	486,117	486,117	484,354	484,354	484,354
Panel E: Bottom Enrollment Quartile						
Bottom Enrollment Quartile	0.010	-0.004	0.023	0.027	0.043	0.034
	(0.024)	(0.024)	(0.025)	(0.027)	(0.028)	(0.024)
Top Enrollment Quartile	0.122***	0.128***	0.057***	0.067**	0.077**	0.004
	(0.023)	(0.025)	(0.016)	(0.022)	(0.025)	(0.022)
Difference	-0.112***	-0.133***	-0.034	-0.040	-0.033	0.030
	(0.033)	(0.035)	(0.030)	(0.035)	(0.038)	(0.033)
N	688,610	688,610	688,610	720,345	720,345	720,345

Teacher-clustered standard errors in parentheses. *p<.05, **p<.01, ***p<.001

Table 8: Association Between GT Status and Teacher Relative Advantage in GT Students

	Math			ELA		
Relative Advantage Definition	(1) >0	(2) >10	(3) >20	(4) >0	(5) >10	(6) >20
GT	0.017***	0.016***	0.009**	0.011*	0.013**	0.010*
	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)
Non-GT High-Testing	0.010**	0.011***	0.006*	0.011**	0.011**	0.008**
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
GT - IA	0.012**	0.009**	0.005	0.010*	0.010**	0.008*
	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
GT - Math	0.015***	0.014***	0.010**	0.009*	0.009*	0.007*
	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.003)
GT - HA	0.016***	0.018***	0.010**	0.008	0.010*	0.006
	(0.005)	(0.004)	(0.003)	(0.005)	(0.004)	(0.004)
GT - ELA	0.005	0.008	0.004	0.010	0.010	0.011
	(0.006)	(0.005)	(0.004)	(0.006)	(0.006)	(0.006)
GT - Arts	0.007	0.005	-0.002	-0.023	-0.018	-0.008
	(0.009)	(0.009)	(0.006)	(0.015)	(0.014)	(0.015)
GT - HG	0.025	0.020	0.012	0.057***	0.060**	0.042*
	(0.017)	(0.014)	(0.015)	(0.014)	(0.020)	(0.017)
N	880,849	880,849	880,849	902,808	902,808	902,808

The outcome is an indicator for relative advantage. >0 (>10; >20) indicates that the GT VAM percentile is more than 0 (10;20) percentile points greater than the low-testing VAM percentile. Teacher-clustered standard errors in parentheses. p<.05, **p<.01, ***p<.001

10 Appendix

Table A.1: Association between Teacher Characteristics and DVAMs by Student Group

	Math DVAM		ELA DVAM	
	(1) High-Testing	(2) GT	(3) High-Testing	(4) GT
National Board Certification	-0.300	1.493	0.910	1.241
	(1.432)	(1.475)	(2.097)	(2.236)
Master's	0.152	-0.562	-0.361	0.495
	(0.665)	(0.778)	(0.968)	(1.015)
Exp: Early Career (3-5 yrs)	1.084	2.369	-1.769	-0.041
	(1.225)	(1.362)	(2.354)	(2.546)
Exp: Middle Career (6-9 yrs)	0.693	0.892	-3.887	-1.312
	(1.325)	(1.541)	(2.497)	(2.733)
Exp: Late Career (10+ yrs)	1.120	1.430	-3.567	-1.152
	(1.384)	(1.617)	(2.669)	(2.939)
Female	1.515	2.406**	1.211	1.841
	(0.802)	(0.910)	(1.025)	(1.044)
Asian	-1.236	-0.989	-1.922	-0.360
	(1.329)	(1.444)	(1.843)	(2.020)
Black	-0.693	-1.755	1.347	1.065
	(1.411)	(1.583)	(1.727)	(2.050)
Latino/a	1.806*	1.707	0.689	1.194
	(0.903)	(0.981)	(1.149)	(1.221)
Other Race	-1.015	-1.578	0.372	-0.721
	(1.624)	(1.938)	(2.737)	(2.845)
N	16,784	16,784	16,504	16,504

School-clustered standard errors are shown in parentheses. All columns include school and year fixed effects and classroom characteristics. Reference groups are novice teachers (1-2 years) and white teachers. *p<.05. **p<.01. ***p<.001.

Table A.2: Association Between GT Status & VAMs, Heterogeneity

Panel A: FRL	Math VAM			ELA VAM		
	High-Testing	GT	Low-Testing	High-Testing	GT	Low-Testing
FRL	-0.012**	-0.012*	-0.015***	-0.005	-0.005	-0.003
	(0.004)	(0.005)	(0.004)	(0.003)	(0.004)	(0.004)
GT	0.054***	0.061***	0.010	0.044***	0.056***	0.003
	(0.013)	(0.016)	(0.009)	(0.013)	(0.015)	(0.013)
FRL × GT	0.004	0.007	0.024*	0.003	-0.003	0.022*
	(0.011)	(0.014)	(0.009)	(0.011)	(0.013)	(0.011)
Sum: GT + FRL×GT	0.059***	0.068***	0.034***	0.048***	0.053***	0.024**
	(0.011)	(0.012)	(0.009)	(0.009)	(0.009)	(0.009)
N	880,849	880,849	880,849	902,808	902,808	902,808

Panel B: EL	Math VAM			ELA VAM		
	High-Testing	GT	Low-Testing	High-Testing	GT	Low-Testing
EL	-0.026**	-0.033***	-0.025**	-0.011	-0.019*	-0.018*
	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)	(0.009)
GT	0.055***	0.063***	0.025**	0.046***	0.053***	0.016
	(0.010)	(0.011)	(0.008)	(0.009)	(0.010)	(0.009)
EL × GT	-0.028	-0.043*	-0.006	-0.046**	-0.060***	0.012
	(0.017)	(0.018)	(0.016)	(0.017)	(0.018)	(0.018)
Sum: GT + EL×GT	0.028	0.020	0.019	-0.000	-0.007	0.028
	(0.017)	(0.019)	(0.016)	(0.016)	(0.017)	(0.017)
N	880,849	880,849	880,849	902,808	902,808	902,808

Panel C: Non-White, Non-Asian	Math VAM			ELA VAM		
	High-Testing	GT	Low-Testing	High-Testing	GT	Low-Testing
Non-White, Non-Asian	-0.020**	-0.019*	-0.018***	-0.019**	-0.017*	-0.014*
	(0.006)	(0.008)	(0.005)	(0.006)	(0.007)	(0.006)
GT	0.057***	0.068***	0.016	0.038**	0.057***	-0.000
	(0.015)	(0.018)	(0.011)	(0.013)	(0.016)	(0.014)
Non-White, Non-Asian × GT	-0.000	-0.004	0.013	0.009	-0.005	0.023
	(0.013)	(0.016)	(0.010)	(0.012)	(0.014)	(0.012)
Sum: GT + nonWA×GT	0.057***	0.064***	0.030***	0.047***	0.052***	0.022*
	(0.010)	(0.011)	(0.008)	(0.009)	(0.009)	(0.009)
N	880,849	880,849	880,849	902,808	902,808	902,808

Panel D: 90%+ FRL School	Math VAM			ELA VAM		
	High-Testing	GT	Low-Testing	High-Testing	GT	Low-Testing
90%+ FRL School	-0.038	-0.011	-0.029	-0.045	-0.031	-0.017
	(0.052)	(0.059)	(0.046)	(0.058)	(0.068)	(0.057)
GT	0.068*	0.055	-0.004	0.014	0.030	-0.044
	(0.031)	(0.035)	(0.022)	(0.028)	(0.031)	(0.030)
90%+ FRL School × GT	-0.006	0.022	0.057*	0.045	0.027	0.095**
	(0.037)	(0.042)	(0.029)	(0.033)	(0.037)	(0.035)
Sum: GT + 90%+ FRL School×GT	0.062**	0.077***	0.053**	0.059**	0.057**	0.051**
	(0.020)	(0.022)	(0.019)	(0.018)	(0.020)	(0.019)
N	486,117	486,117	486,117	484,354	484,354	484,354

Panel E: Bottom Enrollment Quartile	Math VAM			ELA VAM		
	High-Testing	GT	Low-Testing	High-Testing	GT	Low-Testing
Bottom Enrollment Quartile	0.069	0.077	0.078	0.211**	0.216***	0.212**
	(0.059)	(0.058)	(0.064)	(0.069)	(0.065)	(0.070)
GT	0.122***	0.128***	0.057***	0.067**	0.077**	0.004
	(0.023)	(0.025)	(0.016)	(0.022)	(0.025)	(0.022)
Bottom Enrollment Quartile × GT	-0.112***	-0.133***	-0.034	-0.040	-0.033	0.030
	(0.033)	(0.035)	(0.030)	(0.035)	(0.038)	(0.033)
Sum: GT + Bottom Enrollment Quartile×GT	0.010	-0.004	0.023	0.027	0.043	0.034
	(0.024)	(0.024)	(0.025)	(0.027)	(0.028)	(0.024)
N	688,610	688,610	688,610	720,345	720,345	720,345

Teacher-clustered standard errors in parentheses. *p<.05, **p<.01, ***p<.001

Table A.3: Association Between GT Status and Teacher Relative Advantage in High-Testing Students

	Math			ELA		
Relative Advantage Definition	(1) >0	(2) >10	(3) >20	(4) >0	(5) >10	(6) >20
GT	0.019***	0.017***	0.013***	0.010*	0.008	0.011**
	(0.004)	(0.004)	(0.003)	(0.005)	(0.004)	(0.004)
Non-GT High-Testing	0.015***	0.011***	0.008**	0.009**	0.008**	0.009**
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
GT - IA	0.012***	0.012***	0.006*	0.007	0.006	0.008**
	(0.004)	(0.003)	(0.002)	(0.004)	(0.004)	(0.003)
GT - Math	0.018***	0.015***	0.012***	0.011**	0.006	0.007*
	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)
GT - HA	0.018***	0.016***	0.015***	0.005	0.004	0.008*
	(0.005)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)
GT - ELA	0.007	0.006	0.008*	0.013*	0.010	0.012*
	(0.006)	(0.005)	(0.004)	(0.006)	(0.005)	(0.005)
GT - Arts	0.010	0.014	0.004	-0.004	0.002	0.008
	(0.009)	(0.009)	(0.006)	(0.013)	(0.012)	(0.010)
GT - HG	0.026*	0.003	0.019	0.029*	0.044**	0.039*
	(0.013)	(0.018)	(0.012)	(0.012)	(0.016)	(0.016)
N	880,849	880,849	880,849	902,808	902,808	902,808

The outcome is an indicator for relative advantage. >0 (>10; >20) indicates that the high-testing VAM percentile is more than 0 (10;20) percentile points greater than the low-testing VAM percentile. Teacher-clustered standard errors in parentheses. *p<.05, **p<.01, ***p<.001

Table A.4: Association Between GT Status & VAMs Using 10-Student Threshold

Panel A: Overall						
	Math VAM			ELA VAM		
	(1) High-Testing	(2) GT	(3) Low-Testing	(4) High-Testing	(5) GT	(6) Low–Testing
GT	0.059***	0.068***	0.032***	0.050***	0.056***	0.023*
	(0.010)†	(0.011)†	(0.008)	(0.009)†	(0.009)†	(0.009)
Non-GT High-Testing	0.060***	0.066***	0.042***	0.044***	0.047***	0.022**
	(0.009)†	(0.010)†	(0.008)	(0.006)	(0.007)	(0.007)
GT - IA	0.039***	0.045***	0.025***	0.038***	0.043***	0.018*
	(0.008)†	(0.009)†	(0.006)	(0.007)†	(0.008)v	(0.008)
GT - Math	0.053***	0.060***	0.023**	0.037***	0.043***	0.017*
	(0.010)†	(0.011)†	(0.009)	(0.007)†	(0.008)†	(0.007)
GT - HA	0.059***	0.067***	0.031***	0.042***	0.048***	0.021*
	(0.010)†	(0.012)†	(0.008)	(0.009)†	(0.010)†	(0.009)
GT - LA	0.047***	0.054***	0.040***	0.048***	0.046***	0.021
	(0.011)†	(0.012)†	(0.009)	(0.010)†	(0.011)†	(0.012)
GT - Arts	0.014	0.007	0.002	0.011	-0.006	0.005
	(0.016)†	(0.019)	(0.013)	(0.018)†	(0.024)†	(0.018)
GT - HG	0.057*	0.070*	0.022	0.127***	0.162***	0.000
	(0.027)†	(0.033)†	(0.024)	(0.033)†	(0.041)†	(0.030)
N	984,415	984,415	984,415	1,016,467	1,016,467	1,016,467

Panel B: By School Level						
	Math VAM			ELA VAM		
	(1) High-Testing	(2) GT	(3) Low-Testing	(4) High-Testing	(5) GT	(6) Low–Testing
Elementary (3-5)	0.007	0.005	0.009	0.025***	0.024***	0.016**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Middle (6-8)	0.083***	0.096***	0.043***	0.059***	0.068***	0.026*
	(0.014)†	(0.016)†	(0.011)	(0.012)†	(0.013)†	(0.012)
N	984,415	984,415	984,415	1,016,467	1,016,467	1,016,467

This table replicates the analysis from Table 6 using the 10-student threshold for inclusion in analysis. Teacher-clustered standard errors shown in parentheses. All columns include school-grade-year-fixed effects. *p<.05. **p<.01. ***p<.001.
† denotes difference in estimated association is statistically different (p<.001) from the low-testing estimate.

Table A.5: Association Between GT Status & VAMs Using 20-Student Threshold

	Panel A: Overall					
	Math VAM			ELA VAM		
	(1) High-Testing	(2) GT	(3) Low-Testing	(4) High-Testing	(5) GT	(6) Low-Testing
GT	0.059***	0.067***	0.028**	0.046***	0.052***	0.016
	(0.011)†	(0.012)†	(0.009)	(0.009)†	(0.010)†	(0.010)
Non-GT High-Testing	0.063***	0.068***	0.041***	0.043***	0.046***	0.015*
	(0.009)†	(0.011)†	(0.008)	(0.007)†	(0.007)†	(0.007)
GT - IA	0.040***	0.045***	0.021**	0.035***	0.041***	0.012
	(0.008)†	(0.009)†	(0.007)	(0.008)†	(0.009)†	(0.008)
GT - Math	0.053***	0.060***	0.021*	0.034***	0.037***	0.011
	(0.010)†	(0.011)†	(0.009)	(0.008)†	(0.008)†	(0.007)
GT - HA	0.057***	0.065***	0.025**	0.036***	0.044***	0.014
	(0.011)†	(0.012)†	(0.008)	(0.010)†	(0.010)†	(0.010)
GT - LA	0.045***	0.049***	0.039***	0.040***	0.042***	0.016
	(0.011)†	(0.012)†	(0.009)	(0.010)†	(0.012)†	(0.012)
GT - Arts	0.013	0.005	-0.005	0.005	-0.017	0.017
	(0.016)†	(0.019)	(0.012)	(0.020)†	(0.027)†	(0.018)
GT - HG	0.064	0.078	0.017	0.165***	0.188***	0.012
	(0.033)†	(0.040)†	(0.031)	(0.038)†	(0.046)†	(0.026)
N	787,264	787,264	787,264	806,975	806,975	806,975
	Panel B: By School Level					
	Math VAM			**ELA VAM**		
	(1) High-Testing	(2) GT	(3) Low-Testing	(4) High-Testing	(5) GT	(6) Low-Testing
Elementary (3-5)	-0.005	-0.005	-0.009*	0.015**	0.019***	0.011**
	(0.004)	(0.005)	(0.004)	(0.005)	(0.005)	(0.004)
Middle (6-8)	0.078***	0.088***	0.039***	0.053***	0.061***	0.017
	(0.013)†	(0.015)†	(0.011)	(0.012)†	(0.013)†	(0.012)
N	787,264	787,264	787,264	806,975	806,975	806,975

This table replicates the analysis from Table 6 using the 20-student threshold for inclusion in analysis. Teacher-clustered standard errors shown in parentheses. All columns include school-grade-year-fixed effects. *p<.05. **p<.01. ***p<.001. † denotes difference in estimated association is statistically different (p<.001) from the low-testing estimate.

Table A.6: Association Between GT Status & VAMs Using Top/Bottom Tercile Definitions and 15-Student Threshold

	Panel A: Overall					
	Math VAM			**ELA VAM**		
	(1) High-Testing	(2) GT	(3) Low-Testing	(4) High-Testing	(5) GT	(6) Low-Testing
GT	0.054***	0.068***	0.037***	0.046***	0.055***	0.025**
	(0.009)†	(0.011)†	(0.008)	(0.008)†	(0.010)†	(0.009)
Non-GT High-Testing	0.055***	0.065***	0.043***	0.041***	0.046***	0.022***
	(0.008)†	(0.009)†	(0.007)	(0.006)†	(0.006)†	(0.006)
GT - IA	0.037***	0.049***	0.026***	0.033***	0.041***	0.018*
	(0.007)†	(0.009)†	(0.006)	(0.007)†	(0.008)†	(0.007)
GT - Math	0.046***	0.059***	0.028***	0.035***	0.042***	0.020**
	(0.009)†	(0.010)†	(0.008)	(0.007)†	(0.008)†	(0.007)
GT - HA	0.050***	0.062***	0.036***	0.038***	0.044***	0.024**
	(0.009)†	(0.011)†	(0.007)	(0.008)†	(0.010)†	(0.009)
GT - LA	0.044***	0.052***	0.044***	0.037***	0.039***	0.025*
	(0.010)	(0.012)	(0.008)	(0.009)†	(0.011)†	(0.011)
GT - Arts	0.012	0.011	0.005	0.002	-0.020	0.006
	(0.015)	(0.018)	(0.012)	(0.017)†	(0.025)†	(0.019)
GT - HG	0.039	0.058	-0.001	0.163***	0.215***	-0.030
	(0.024)†	(0.031)†	(0.023)	(0.039)†	(0.054)†	(0.032)
N	976,645	976,645	976,645	1,004,961	1,004,961	1,004,961

	Panel B: By School Level					
	Math VAM			**ELA VAM**		
	(1) High-Testing	(2) GT	(3) Low-Testing	(4) High-Testing	(5) GT	(6) Low-Testing
Elementary (3-5)	-0.002	-0.002	0.004	0.021***	0.021***	0.021***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Middle (6-8)	0.078***	0.100***	0.051***	0.055***	0.068***	0.026*
	(0.013)†	(0.016)†	(0.011)	(0.011)†	(0.013)†	(0.012)
N	976,645	976,645	976,645	1,004,961	1,004,961	1,004,961

This table replicates the analysis from Table 6 using top and bottom tercile definition of high- and low-testing. Teacher-clustered standard errors shown in parentheses. All columns include school-grade-year-fixed effects. *p<.05. **p<.01. ***p<.001.
† denotes difference in estimated association is statistically different (p<.001) from the low-testing estimate.

Table A.7: Association Between GT Status & VAMs Using Top/Bottom Quintile Definitions and 15-Student Threshold

	Panel A: Overall					
	Math VAM			**ELA VAM**		
	(1) High-Testing	(2) GT	(3) Low-Testing	(4) High-Testing	(5) GT	(6) Low-Testing
GT	0.061***	0.065***	0.027**	0.050***	0.056***	0.006
	(0.011)†	(0.012)†	(0.008)	(0.009)†	(0.010)†	(0.009)
Non-GT High-Testing	0.065***	0.068***	0.037***	0.047***	0.048***	0.009
	(0.010)†	(0.011)†	(0.008)	(0.007)†	(0.008)†	(0.007)
GT - IA	0.043***	0.046***	0.023***	0.041***	0.044***	0.004
	(0.008)†	(0.009)†	(0.007)	(0.008)†	(0.009)†	(0.008)
GT - Math	0.053***	0.056***	0.017*	0.031***	0.039***	0.006
	(0.010)†	(0.011)†	(0.009)	(0.008)†	(0.008)†	(0.008)
GT - HA	0.058***	0.063***	0.027**	0.043***	0.048***	0.004
	(0.012)†	(0.013)†	(0.008)	(0.010)†	(0.011)†	(0.009)
GT - LA	0.044***	0.048***	0.037***	0.046***	0.045***	0.012
	(0.012)†	(0.013)†	(0.009)	(0.011)†	(0.012)†	(0.012)
GT - Arts	0.005	0.005	-0.006	0.005	-0.020	0.022
	(0.018)	(0.020)	(0.015)	(0.023)†	(0.028)†	(0.022)
GT - HG	0.072	0.083	0.021	0.195***	0.210***	0.002
	(0.037)†	(0.042)†	(0.030)	(0.043)†	(0.049)†	(0.029)
N	783,677	783,677	783,677	797,170	797,170	797,170
	Panel B: By School Level					
	Math VAM			**ELA VAM**		
	(1) High-Testing	(2) GT	(3) Low-Testing	(4) High-Testing	(5) GT	(6) Low-Testing
Elementary (3-5)	0.002	0.001	-0.002	0.020***	0.017**	0.012**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Middle (6-8)	0.078***	0.084***	0.036***	0.057***	0.065***	0.005
	(0.014)†	(0.016)†	(0.011)	(0.012)†	(0.013)†	(0.012)
N	783,677	783,677	783,677	797,170	797,170	797,170

This table replicates the analysis from Table 6 using top and bottom quintile definition of high- and low-testing. Teacher-clustered standard errors shown in parentheses. All columns include school-grade-year-fixed effects. *p<.05. **p<.01. ***p<.001.
† denotes difference in estimated association is statistically different (p<.001) from the low-testing estimate.

References

Ahn, T., E. Aucejo, and J. James (2025). The Importance of Student-Teacher Matching: A Multidimensional Value-Added Approach. *Conditionally accepted, Review of Economics and Statistics*.

Bangel, N. J., S. M. Moon, and B. M. Capobianco (2010). Preservice Teachers' Perceptions and Experiences in a Gifted Education Training Model. *Gifted Child Quarterly 54*(3), 209–221.

Barbier, K., E. Struyf, K. Verschueren, and V. Donche (2023). Fostering Cognitive and Affective-Motivational Learning Outcomes for High-Ability Students in Mixed-Ability Elementary Classrooms: A Systematic Review. *European Journal of Psychology of Education 38*(1), 83–107.

Berman, K. M., R. A. Schultz, and C. L. Weber (2012). A Lack of Awareness and Emphasis in Preservice Teacher Training: Preconceived Beliefs About the Gifted and Talented. *Gifted Child Today 35*(1), 18–26.

Biasi, B., C. Fu, and J. Stromme (2021). Equilibrium in the Market for Public School Teachers: District Wage Strategies and Teacher Comparative Advantage. Technical report, National Bureau of Economic Research.

Bishop, W. E. (1968). Successful Teachers of the Gifted. *Exceptional Children 34*(5), 317–325.

Chan, D. W. (2001). Characteristics and Competencies of Teachers of Gifted Learners: The Hong Kong Teacher Perspective. *Roeper Review 23*(4), 197–202.

Chessman, A. M. (2010). *Teacher Attitudes and Effective Teaching Practices for Gifted Students at Stage 6*. Ph. D. thesis, University of New South Wales.

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

Chingos, M. and P. Peterson (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.

Dai, D. Y. (2013). Excellence at the Cost of Social Justice? Negotiating and Balancing Priorities in Gifted Education. *Roeper Review 35*(2), 93–101.

David, H. (2011). The Importance of Teachers' Attitude in Nurturing and Educating Gifted Children. *Gifted and Talented International 26*(1-2), 71–80.

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

Dixon, F. A., N. Yssel, J. M. McConnell, and T. Hardin (2014). Differentiated Instruction, Professional Development, and Teacher Efficacy. *Journal for the Education of the Gifted 37*(2), 111–127.

Egalite, A., B. Kisida, and M. Winters (2015). Representation in the Classroom: The Effect of Own-race Teachers on Student Achievement. *Economics of Education Review 45*, 44–52.

Eilam, B. and H. E. Vidergor (2011). Gifted Israeli Students' Perceptions of Teachers' Desired Characteristics: A Case of Cultural Orientation. *Roeper Review 33*(2), 86–96.

Fraser-Seeto, K. T., S. Howard, and S. Woodcock (2015). An Investigation of Teachers' Awareness and Willingness to Engage with a Self-Directed Professional Development Package on Gifted and Talented Education. *Australian Journal of Teacher Education 40*(40), 1–14.

Gallagher, J. (2015). Political Issues in Gifted Education. *Journal for the Education of the Gifted 38*(1), 77–89.

Geake, J. G. and M. Gross (2008). Teachers' Negative Affect Toward Academically Gifted Students: An Evolutionary Psychological Study. *Gifted Child Quarterly 52*(3), 217–231.

Grissom, J. and C. Redding (2016). Discretion and Disproportionality: Explaining the Underrepresentation of High-Achieving Students of Color in Gifted Programs. *AERA Open 2*(1), 1–25.

Gross, M. (1994). Changing Teacher Attitudes to Gifted Students through Inservice Training. *Gifted and Talented International 9*(1), 15–21.

Hansen, J. B. and J. F. Feldhusen (1994). Comparison of Trained and Untrained Teachers of Gifted Students. *Gifted Child Quarterly 38*(3), 115–121.

Hanushek, E. and S. Rivkin (2010). Generalizations About Using Value-Added Measures of Teacher Quality. *American Economic Review 100*(2), 267–271.

Joshi, E., S. Doan, and M. Springer (2018). Student-Teacher Race Congruence: New Evidence and Insight from Tennessee. *AERA Open 4*(4), 1–25.

Kane, T. and D. Staiger (2008). Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation. Technical Report No. w14607, National Bureau of Economic Research.

Kettler, T., J. Russell, and J. Puryear (2015). Inequitable Access to Gifted Education: Variance in Funding and Staffing Based on Locale and Contextual School Variables. *Journal for the Education of the Gifted 38*(2), 99–117.

Khalil, M. and Z. Accariya (2016). Identifying "Good" Teachers for Gifted Students. *Creative Education 07*(03), 407–418.

Klassen, R. and V. Tze (2014). Teachers' Self-Efficacy, Personality, and Teaching Effectiveness: A Meta-Analysis. *Educational Research Review 12*, 59–76.

Lassig, C. (2015). Teachers' Attitudes Towards the Gifted: The Importance of Professional Development and School Culture. *Australasian Journal of Gifted Education 24*(2), 32–42.

LAUSD (2024). Gifted/Talented Programs. Accessed on 07/22/24.

Lockwood, J. and D. McCaffrey (2014). Correcting for Test Score Measurement Error in ANCOVA Models for Estimating Treatment Effects. *Journal of Educational and Behavioral Statistics 39*(1), 22–52.

Loeb, S., J. Soland, and L. Fox (2014). Is a Good Teacher a Good Teacher for All? Comparing Value-Added of Teachers With Their English Learners and Non-English Learners. *Educational Evaluation and Policy Analysis 36*(4), 457–475.

Maddux, C. D., I. Samples-Lachmann, and R. E. Cummings (1985). Preferences of Gifted Students for Selected Teacher Characteristics. *Gifted Child Quarterly 29*(4), 160–163.

Master, B., S. Loeb, C. Whitney, and J. Wyckoff (2016). Different Skills? Identifying Differentially Effective Teachers of English Language Learners. *The Elementary School Journal 117*(2), 261–284.

Matheis, S., L. Kronborg, M. Schmitt, and F. Preckel (2017). Threat or Challenge? Teacher Beliefs About Gifted Students and Their Relationship to Teacher Motivation. *Gifted and Talented International 32*(2), 134–160.

Mathijssen, A. C. S., L. Hoogeveen, and E. Jen (2021). Professional Development in Gifted Education: Training Teachers to Meet the Needs of Gifted Students. In *Encyclopedia of Teacher Education*, pp. 1–5. Singapore: Springer Singapore.

Mills, C. J. (2003). Characteristics of Effective Teachers of Gifted Students: Teacher Background and Personality Styles of Students. *Gifted Child Quarterly 47*(4), 272–281.

NAGC (2014a). Differentiating Curriculum and Instruction for Gifted and Talented Students. Technical report, National Association for Gifted Children.

NAGC (2014b). Knowledge and Skill Standard in Gifted and Talented Education for All Teachers. Technical report, National Association for Gifted Children.

National Association for Gifted Children (2005). The History of Gifted and Talented Education. Accessed on 07/29/2024.

Olszewski-Kubilius, P. and S. Steenbergen-Hu (2017). Blending Research-Based Practices and Practice-Embedded Research: Project Excite Closes Achievement and Excellence Gaps for Underrepresented Gifted Minority Students. *Gifted Child Quarterly 61*(3), 202–209.

Perković Krijan, I. and E. Borić (2014). Teachers' Attitudes Towards Gifted Students and Differences in Attitudes Regarding the Years of Teaching / Stavovi učitelja prema darovitim učenicima i razlike u stavovima s obzirom na staž. *Croatian Journal of Education - Hrvatski časopis za odgoj i obrazovanje 17*, 165–178.

Peters, S. and K. Engerrand (2016). Equity and Excellence: Proactive Efforts in the Identification of Underrepresented Students for Gifted and Talented Services. *Gifted Child Quarterly 60*(3), 159–171.

Podolsky, A., T. Kini, and L. Darling-Hammond (2019). Does Teaching Experience Increase Teacher Effectiveness? A Review of US Research. *Journal of Professional Capital and Community 4*(4), 286–308.

Rinn, A., R. Mun, and J. Hodges (2020). 2018-2019 State of the States in Gifted Education. Technical report.

Sahin, F. and F. Levent (2015). Examining the Methods and Strategies which Classroom Teachers use in the Education of Gifted Students. *The Online Journal of New Horizons in Education 5*(3), 73–82.

Stephens, K. R. (2009). Can Empathy for Gifted Students Be Nurtured in Teachers? *Gifted Children 3*, 1–5.

Tischler, K. and W. J. Vialle (2009). Gifted Students' Perceptions of the Characteristics of Effective Teachers. Technical report, University of Wollongong.

Troxclair, D. A. (2013). Preservice Teacher Attitudes Toward Giftedness. *Roeper Review 35*(1), 58–64.

U.S. Department of Education (2017). Digest of Education Statistics. Technical report, U.S. Department of Education.

Wood, W. and I. Lai (2022). *The Effect of Same-race Teachers and Faculty on Student Test Scores.* Doctoral Dissertation, Michigan State University.

Wood, W. J., I. Lai, N. R. Filosa, S. A. Imberman, N. D. Jones, and K. O. Strunk (2024). Are Effective Teachers for Students With Disabilities Effective Teachers for All?

Woods, J. W. (2004). *The Characteristics of Successful and Effective Teachers of the Gifted.* Ph. D. thesis, University of Tennessee, Knoxville.