Second Time's the Charm? How Repeat Student-Teacher Matches Build Academic and Behavioral Skills

Leigh Wedenoja
Rockefeller Institute of Government

John P. Papay
Brown University

Matthew A. Kraft
Brown University

July 21, 2020

Abstract

We examine the dynamic nature of student-teacher match quality by studying the effect of having a teacher for more than one year. Using state-wide data from Tennessee and panel methods, we find that having a repeat teacher improves achievement and decreases absences, truancy, and suspensions. These results are robust to a range of tests for teacher and student sorting. White girls benefit most academically from repeat teachers and boys of color benefit most behaviorally. Effects increase with the share of repeat students in a teacher's class suggesting that intentional classroom assignments policies such as looping may have even larger benefits.

1 Introduction

Staffing all classrooms with effective teachers is a perennial goal of education policy. This focus is well justified given the large impacts teachers have on students' test scores (Rivkin et al., 2005; Rockoff, 2004), as well as their socioemotional skills (Jackson, 2018; Cunha & Heckman, 2008; Heckman et al., 2006; Gershenson, 2016; Kraft, 2019; Blazar & Kraft, 2017; Ladd & Sorensen, 2017) and long-run outcomes (Chetty et al., 2014; Jackson, 2018). Researchers and policymakers often frame this challenge in a way that assumes teacher effectiveness as a fixed characteristic; a teacher is either effective or not. Policy solutions gravitate towards human capital strategies to recruit, select, and retain high-quality teachers and dismiss low-performing teachers. However, recent studies suggest that teacher effectiveness is both dynamic and context-dependent – it evolves over time (Rockoff, 2004; Harris & Sass, 2011; Ost, 2014; Papay & Kraft, 2015) and it depends on the match between a teacher, her school, and her students (Jackson, 2013).

We provide further evidence on the importance of match quality for teacher effectiveness and the potential to improve the productivity of the current teacher workforce through student-teacher matching. Prior research on student-teacher match has largely viewed match quality as fixed, focusing on the benefits to matching based on characteristics such as shared demographics (Dee, 2005; Clotfelter, Ladd & Vigdor, 2006; Jackson, 2009; Egalite, Kisidab & Winters, 2015; Gershenson et al., 2018). We explore the dynamic nature of these student-teacher matches. Over time, teachers develop relationships with students and learn how to tailor their instruction to students' individual learning styles. The education sociology and psychology literatures describe student-teacher relationships as dynamic and continuously developing over time. These studies document that stronger student-teacher relationships are associated with higher academic achievement, fewer disciplinary problems, and decreased risky behavior (Hamre & Pianta, 2006; Eisenhower & Baker, 2007, DiLalla et al.,

2004, Hamre & Pianta, 2001).

We examine the dynamic nature of student-teacher match quality by studying students who are taught by the same teacher for more than one year. Educators and parents have long discussed the potential positive effects of repeat student-teacher interactions most commonly in terms of intentionally looped classes. Looping is a policy in which a teacher stays with the same classroom of students for two or more years (Burke, 1997; Franz et al., 2010; George & Lounsbury, 2000). Advocates of looping posit that, because they spend longer with students, teachers in looped classrooms build stronger relationships with their students and can flexibly adjust the curriculum over two years to better utilize classroom time.

However, formally looped classrooms are not the most common form of repeated student-teacher interaction. In our data, less than 5% of teachers with repeat students could be considered "looping" teachers. Instead, teachers most often encounter repeat students when they move grades within their school or teach the same subject across multiple grades. These unsystematic repeat student-teacher matches are relatively common, 44% of the students in our data have at least one repeat teacher in math or ELA between 3rd through 11th grade. While unsystematic forms of repeat student-teacher matches may not reflect the full potential benefit of intentional looping, students and teachers may still benefit from a second year together if spending more time with a teacher improves student-teacher relationships.

Our analyses leverage panel data on students and teachers from Tennessee to estimate the effect of repeated student-teacher matches on outcomes for students in grades 3 through 11. The breadth of the data allows us to examine the differential effect of repeat teachers on achievement across a student's career using panel methods with high-dimensional fixed effects. Our estimates derive from teachers (a) moving across grades or (b) teaching more than one grade in a given year.[1] We compare outcomes in years when students are in a classroom with a teacher they have had

[1]We explicitly exclude repeat interactions that occur when students are retained in a grade.

before (a repeat teacher) to years in which the student has a teacher for the first time. We account for sorting of students to repeat teachers in two complementary ways: by making within-student comparisons with student fixed effects, and controlling flexibly for lagged outcomes in education production function-style models. We conduct a range of additional tests to rule out sorting as a threat to validity.

We find that repeat teachers increase test scores for students at all grade levels, similar to earlier results for elementary school students by Hill and Jones (2018). We also find that these repeat interactions decrease disciplinary infractions for students across grade levels, and improve attendance in high school by reducing truancy. These results suggest that teachers do indeed become more effective working with students in their second year together, providing evidence that match quality is dynamic and can improve over time.

We also find substantial heterogeneity in the effect of repeat teachers. Positive test score gains are concentrated among higher-performing students and white female students. However, improvements in attendance and discipline are driven primarily by male students of color as well as students who have lower test-scores. We find less heterogeneity by teacher characteristics: students benefit from repeating with teachers of all levels of experience, and there is little evidence that repeating with a teacher who has low value added magnifies the negative effect of that teacher. Finally, we find evidence of positive spillovers, as both repeat and non-repeat students benefit in classes with a large share of repeat students.

2 Temporal Dynamics of Student-Teacher Match

Standard education production function models apply the logic of firm production to schools, asserting that a student's academic and behavioral outcomes in a given year are the result of the student's ability, effort, and all other productive inputs including teacher, school, and parent influences the student receives in the current year and

from all past years over the course of her career (Todd & Wolpin, 2003). Recent research has developed in detail some elements of this model, such as highlighting the critical role of teacher inputs as within-school productive factors, documenting the substantial variation in productivity across teachers, and exploring the dynamic nature of teacher productivity over time.

These production function models have also incorporated a range of other theoretical advances from literature on the firm; most relevant here are the canonical models of worker-firm match in determining productivity (Jovanovic, 1979). Jackson (2013) documented the presence of teacher-school match effects, separating intuitively the main effects of teacher human capital and school production technology on student outcomes from the differential impact on productivity from the teacher-school match. In other words:

$$y_{ijst} = f(y_{i,t-1}) + \delta_s + \tau_j + \gamma_{js} + \eta_{ijst} \tag{1}$$

Here, student outcomes in given year for student i with teacher j in school s at time t depend not only on prior family and school investments in the student's learning, proxied by $f(y_{i,t-1})$, but also contemporaneous learning derived from school productivity (δ_s), teacher productivity (τ_j), and teacher-school match quality (γ_{js}).

These models focus on match quality between teachers and schools (workers and firms). However, the concept of match quality can easily be extended to the match between a teacher and her students. There is a long literature documenting the presence of such matches, for example documenting that students of color benefit from being taught by a teacher of the same race, independent of that teacher's overall level of productivity (Dee, 2005; Clotfelter, Ladd & Vigdor, 2006; Jackson, 2009; Egalite, Kisidab & Winters, 2015; Gershenson et al., 2016) and that some teachers are more effective with English learners than non-English learners, and vice versa (Loeb, Soland, & Fox, 2014). These models implicitly hypothesize the presence of a match parameter between students and teachers (λ_{ij}), such that individual teachers

are more productive with certain types of students, or individual students exert more effort or learn better from certain types of teachers.

But match quality, whether teacher-school or student-teacher, need not be fixed over time. How experienced a teacher is with a particular student is an important temporal component of the teacher's effect because increased experience with a given student is likely to improve match quality as teachers build student-specific human capital. We can envision a student-teacher match quality as dynamic such that the match quality parameter is time varying (λ_{ijt}). Our analyses focus on estimating the relationship between this dynamic student-teacher match quality input in the education production function and student outcomes as implied by the production function:

$$y_{ijst} = f(y_{i,t-1}) + \delta_s + \tau_j + \gamma_{js} + \lambda_{ijt} + \eta_{ijst} \tag{2}$$

This function explicitly allows the student-teacher match to change over time. If match improves when teachers and students are more exposed to each other then we would expect $\lambda_{ijt} > \lambda_{ijt-1}$.

3 Data and Descriptive Analysis

Data for this paper come from the Tennessee Education Research Alliance and are compiled from Tennessee Department of Education administrative records. Tennessee links all students to teachers in tested grades and subjects. We identify each teacher that a student has in a tested grade and subject and whether the student has had that teacher before. We focus on students in Math and English Language Arts (ELA) who take grade-level Tennessee Comprehensive Assessment Program (TCAP) tests in grades 3 – 8 and specific-subject tests in high school for Algebra I, Algebra II, Geometry, English I, English II, and English III.[2] These courses cover nearly all students in grades 3 – 8 and the vast majority of high school students in grades 9

[2]Given the availability of high-school testing data, we focus on math and ELA. Results including Science and Social Studies exams are quite similar.

– 11. We use elementary and middle school data for spring 2007 – spring 2015 and high school data for spring 2007 – spring 2017.[3]

We focus our analyses on students' primary teachers for each grade and subject.[4] We further exclude students who are held back in their grade or who are taking a test for the second time. While these students may well have the same teacher for a second year, there are many other reasons why we would expect their test performance to be better when repeating the same content. Our final analytical sample includes 1.3 million unique students and more than 50,000 unique teachers. Additional student data include demographic information on gender, race, ethnicity, special education status, limited English proficiency, and receipt of free or reduced-price lunch.

Beyond test scores, we also examine observed behaviors which are important outcomes in themselves and serve as proxies for un-observed socioemotional skills, sometimes referred to as "non-cognitive" skills (Lleras 2008; Bertrand & Pan 2013; Kautz and Zanoni 2014; Heckman et al., 2016; Jackson, 2018). The behaviors we examine – attendance, truancy, and discipline – are all associated with well-established psychometric measures of socioemotional skills (Heckman, 2006; John et al., 1994; Barbaranelli et al., 2003; Lounsbury et al., 2004; Carneiro, Crawford & Goodman 2007; Duckworth et al., 2007). Our data on attendance and suspension are at the student-school-year level and are not tied to a specific teacher or class period. We use two measures of attendance: total days absent during the school year and total days with unexcused (truant) absences for high school students.[5] Total absences account for the amount of time that a student is out of school, whereas truant absences in

[3]Due to challenges with TCAP testing in spring 2016 and 2017, we omit those years from the analysis. Not all EOC test subjects are available for all data years: Algebra I, English I, and English II are available for all years of the data; Algebra II and English III were introduced in 2012; and Geometry was introduced in 2016.

[4]Because we are interested in the effect of having a repeated primary classroom teacher, rather than a repeated support teacher or ELL specialist, we limit the sample to student-teacher matches where the teacher claims at least a 50% responsibility for the student in a given subject, the student has been enrolled in Tennessee schools for at least half the school year, the teacher's job title is "classroom teacher," and the teacher has at least 5 students and no more than 200 students in the subject. These restrictions eliminate less than 1% of the sample.

[5]The distinction between excused and un-excused absences is imprecise and complicated in elementary and middle school when parents are primarily responsible for their child's attendance. Truancy is well defined and linked to school engagement in high school.

high school can be used as a measure of student engagement in school (Wedenoja, 2017; Imberman, 2011). Suspension is measured on the extensive margin as whether or not a student has had a serious disciplinary incident resulting in an in-school or out-of-school suspension during the school year. Because of evidence about racial bias in school discipline, we control for student race in all models where we examine these outcomes (Owens & McLanahan, 2020).

Table 1 contains the average demographic characteristics of students and teachers in our sample and documents differences between students who have at least one repeat teacher and students who have no repeat teachers. Students with at least one repeat teacher are more likely to be white and low income and have lower test scores in the year before they are repeat students, but are also less likely to be absent or truant. These differences largely reflect the fact that repeat student-teacher interactions are more frequent in smaller schools, which tend to be whiter and lower-performing. We see no differences if we compare students in the same school.

Repeat students also differ in the teachers they are likely to encounter in their classrooms. We compare teachers for whom more than 10% of their students are students they have taught before to teachers with less than 10% repeat students. Overall teachers with at least 10% repeat students have an extra year of experience, are a year older, and, much like their students, are 2 percentage points more likely to be white. Despite the apparent higher qualifications of repeat teachers, they have a slightly lower measured effectiveness (VAM) compared to their peers.

3.1 Prevalence of Repeat Student-Teacher Assignments in Tennessee

We define a student-teacher pair as a "repeat" in year t if the teacher is a student's primary teacher in a subject in both year t and a previous year, and the student is not repeating a grade or test. Teacher repetition is not limited to consecutive years: an eighth-grade student with the same math teacher she had in sixth grade has a

repeat teacher. Repetition in a given year is relatively rare. Only 6% of students have a repeat teacher in math or ELA at any one time. However, repetition is common over the course of a student's career: 44% of the students we observe from grades 3 – 11 have had at least one repeat teacher over the course of their school career. These repetition estimates are likely to understate the true percentage of students who have repeat teachers in any one year and especially the percentage who have ever had a repeat teacher because we only focus on math and ELA teachers in tested grades and subjects.[6]

There is substantial variation in the prevalence of repeat student-teacher assignments by grade and subject. In Table 2, we show the number of student-year observations in each tested subject as well as the probability we observe a repeat teacher for that student in that subject. Not surprisingly, repeat teachers are more common in higher grades ranging from only 2% of fourth grade students to 11% of eighth grade students in each subject.

Teachers are much more likely to have repeat students than students are to have repeat teachers. This is partially mechanical as a student only has one teacher per subject whereas a teacher has many students. Column 5 of Table 2 gives the percentage of teachers in each tested subject that have any repeat students and Column 6 gives the percentage of teachers that have at least 10% repeat students.[7] Again, repeat matches are most common in 8th grade when 33% of teachers have at least one repeat student and nearly 30% have at least 10% repeat students. Few of these teachers are engaged in formal looping where they are teaching intact classes of students in multiple years – Column 7 shows that, in any given year, only 1.5% of teachers are assigned to classrooms where at least 90% of their students are repeat

[6]Due to our reliance on tested subjects, we are particularly likely to understate the prevalence of repetition in elementary grades and in high school as most courses do not have EOC exams and as a result do not have student-teacher links. If a student had their math teacher previously in an un-tested subject, such as statistics, we would not identify that repetition. EOC coverage in English is much better. Virtually all students in our sample take both the English I and English II EOC exams during high school, and a little more than half take English III as well. As such, English III has the highest percentage of students repeating with a teacher (9.5%, compared to approximately 5% for the other EOC subjects).

[7]Elsewhere in the paper we use the 10% threshold as the cutoff for being a "repeat teacher.

students.

3.2 Sources of Repeat Student-Teacher Assignments

We examine the sources of repeat assignments in a sample of repeat teachers who have had at least 10% of their students before and whom we observe the year before and after they are repeat teachers. Table 3 documents the paths through which these teachers end up with repeat students across different grades and subjects. The lion's share of repeat student-teacher assignments that we observe arise from unintentional looping – teachers moving grades or consistently teaching in multiple grades.

In elementary school (grades 4 and 5) most teachers have repeat students because they moved from a lower grade to a higher grade. This type of repetition accounts for 87% of fourth grade repeat teachers and 67% of fifth grade repeat teachers. These moves into higher grades also tend to be permanent moves rather than a temporary reassignment or an intentional loop. Over two thirds of these teachers stay in that same grade the following year. The remaining repeat elementary school teachers are those who teach in multiple grades simultaneously.

In middle and high school, the story is different. In these grades (6 – 11) most repetition is due to teachers having classrooms in multiple grades at the same time. Here, 53% to 77% of all repeat teachers have students in multiple grades. Furthermore, the vast majority of repeat teachers after grade 6 are teaching in the same grade they taught the year before and continue to teach in that grade the following year.

4 Econometric Methodology and Identification

We explore the dynamic nature of student-teacher match by estimating whether student outcomes improve more in years when they have a repeat assignment to a teacher than in other years. While relatively few students are assigned to repeat teachers as

part of an intentional looping policy, other endogenous classroom assignment policies may affect which students experience repeat teachers in which years. This potential endogeneity makes it inappropriate to directly compare the outcomes of students with repeat teachers to those without repeat teachers even if the repetition is not intentional. In the absence of random assignment of students to repeat teachers, we follow the teacher effects literature and control for non-random assignment of students to teachers and classrooms to address concerns over classroom assignment and other forms of endogeneity.

We employ two complimentary identification strategies: controlling for student fixed effects and controlling directly and flexibly for lagged outcomes. Because we believe that student sorting into classrooms is the largest threat to validity, our preferred specification controls for student fixed effects. Thus, our identification derives from explicitly comparing students in years and subjects in which they have repeat teachers to themselves in years in which they do not. We present complementary results controlling for student outcomes from the prior year– a cubic of math and reading test scores, log of attendance, and an indicator for whether the student was suspended. This lagged specification complements our primary identification strategy by comparing different students with similar outcome histories in the same year, district, and subject.

Across both specifications, we also control for district-by-year fixed effects, test-by-year fixed effects, and time varying school-level and class-level demographics. The district-by-year fixed effects control for yearly differences in school calendars, leadership, testing policies, and other idiosyncrasies that vary between districts and years. The school-level and class-level demographic and lagged outcome averages account for peer effects and the sorting of students between and within schools and within districts.

Specially, we fit the following models:

$$y_{ijlt} = \beta_0 + \beta_r * REP_{ijlt} + \beta_e^{'} S_{st} + \beta_c^{'} C_{jlt} + \beta_p^{'} P_{jt} + \beta_x^{'} X_{it} + \alpha_i + \delta_{dt} + \epsilon_{ijlst} \quad (3)$$

$$y_{ijlt} = \beta_0 + \beta_r * REP_{ijlt} + \beta_e^{'} S_{st} + \beta_c^{'} C_{jlt} + \beta_p^{'} P_{jt} + \beta_x^{'} X_{it} + \beta_y^{'} Y_{it-1} + \delta_{dt} + \epsilon_{ijlst} \quad (4)$$

Where y_{ijlt} is the outcome of interest for student i with teacher j in subject l and year t. In both models $REPEAT$ is a dummy variable such that $REP = 1$ when a teacher has had the student before and 0 otherwise and β_r measures the effect of being with a teacher for a second (or more) time. As discussed above, X is a vector of student control variables, S_{st} is a vector of school-by-year control variables for school s in year t, C_{jlt} is a vector of classroom controls for teacher j in subject l in year t, and δ_{dt} is the district by year fixed effect. Equation 3, also contains a fixed effect, α_i for student i and Equation 4 contains a vector of lagged outcomes, Y_{it-1}, including a cubic in lagged test scores, for both subjects, as well as lagged absences and suspension. Because teachers improve their effectiveness as they gain experience, particularly early in their career (Rockoff, 2004; Harris & Sass, 2011; Papay & Kraft, 2015) and teachers mechanically have more experience during the second year with a student compared to the first, we include a full set of dummy variables (vector P_{jt}) for teacher experience in all of our models.

The definition of control variables and lagged outcomes varies slightly in high school. Because students take math classes in different sequences and different years of high school, we use eighth grade test scores in lieu of lagged high-school test scores. This limits our high school sample to students who also attended eighth grade in Tennessee, but has the added benefit of not controlling for lagged scores that may have also been impacted by the repeat teacher. Further, following Jackson (2014), we control for "track" in high school; all students in the same grade take the same ELA EOC exam, but we control for students who take the Algebra I EOC exam before ninth grade.

Our specification for behavioral outcomes closely mirrors our achievement models. Because behavioral outcomes are measured at the student-year rather than student-teacher-year level, we attribute these outcomes to all the student's observed teachers. In elementary school most students have a single main teacher so there is no effective difference in our models. However, for students in middle and high school who have multiple teachers, we weight the student's absence and suspension by the number of different teachers. Furthermore, both attendance and truancy have heavily skewed distributions with long right-hand tails. To account for this skewness and the fact that many high school students have no truancies, we transform the absence and truancy outcomes to $ln(totalabsences + 1)$ and $ln(totaltruancies + 1)$. We measures suspension on the extensive margin as a binary variable for whether or not the student was suspended (in-school or out of school) during the year. This abstracts from different school policies relating to severity of suspensions and other, lesser disciplinary consequences that may be recorded differently across schools.

Our measures of teachers' effects on student behavior are by necessity conservative and should be viewed as lower bounds of the true effect of repeat teachers on behavioral outcomes for two reasons. First, we do not observe class-specific attendance or discipline; if a teacher reduces absenteeism within only her class, these effects will be diluted across several teachers. Second, full day absences likely understate the full impact as most missed absences are for single classes rather than full days (Liu & Loeb, 2019). Similarly, we measure suspension on the extensive margin and only estimate the effect of repeat teachers on the probability that the student is ever suspended during the year, not on whether a repeat teacher reduces the frequency or length of suspensions or impacts other, lesser disciplinary consequences.

5 Results

Having a repeat teacher improves student test scores. Panel A of Table 4 contains pooled results for Math and Reading/ELA in row one and Math and Reading/ELA separately in rows two and three for both the student fixed effects and lagged outcome specifications from Equation 2. Overall, having a repeat teacher is associated with a 0.02 SD increase in test scores, which represents an improvement in teacher productivity of 10-15% in a given year.[8] Estimates of the effect of repeat teachers are quite consistent across school levels, subjects, and model specifications.

In Panel B we present similar results for the three behavioral outcomes – absences, suspensions, and truancies. Having a repeat teacher reduces absences and suspensions across all grade levels and specifications. The reduction of absences in high school can be almost entirely attributed to a reduction in truancies. Despite the consistent and significant reduction in absences and truancies, the estimated effects are themselves modest. Absences are reduced by 0.5% overall and the probability students are suspended during the year is reduced by 1 percentage point (a 10% reduction). There is no significant difference across grade levels in the effect on suspensions and absences, however the reduction in high school truancies is likely more meaningful than the effect on absences in earlier grades because even a single truancy in 9th grade predicts lower test scores, a lower GPA, and a reduced probability of graduating from high school (Wedenoja, 2018). Conversely the percentage reduction in disciplinary incidence is higher in elementary and middle school because of the lower levels of suspensions overall compared to high school.

5.1 Robustness

Our preferred student-fixed-effects models identify the effect of repeat teachers by comparing students who have a repeat teacher to themselves in years in which they

[8] Despite similar point estimates across school levels, effects in high school are functionally larger and roughly equivalent to an additional month of learning (Hill et al., 2008). The results for elementary and middle school, by comparison, roughly correspond to an additional week or two of learning.

do not. This approach is attractive because it accounts for the sorting of students to repeat classrooms based on observable and unobservable time-invariant student characteristics. However, other types of selection may be at play. For example, it may be that more effective teachers or more effective schools are those where repeat student-teacher interactions are most likely. Descriptively, we find that this is not the case: the average value-added (conditional on experience) of repeat teachers is lower than non-repeat teachers (Table 2). Repeat students have teachers with measured value added 0.03 SD lower than teachers of non-repeat students and that difference is consistent when controlling for school FE, school-by-year FE, and test FE as well as teacher experience.

To further explore these potential sources of sorting, we restrict our analyses in several ways to account for differences in policies and characteristics between schools and classrooms. First, we replace our district fixed effects with school-by-year fixed effects, identifying differences in achievement of students with repeat teachers and those without in the same school and year. We further limit this analysis by fitting models with teacher-by-year fixed effects, explicitly identifying the effect of the repeat teacher by comparing the effect of that teacher to students in the same classroom who did and did not have the teacher in the past. The effective variation used to identify these effects comes from a somewhat different sample than our preferred approach, because schools or classrooms with only repeat students, or only non-repeat students, do not contribute to the estimates.

In Table 5, we compare our preferred estimates to these alternative specifications. In general, our results are robust to the inclusion of school-by-year and teacher-by-year fixed effects. We present both our student fixed effects and lagged specifications with school-by-year instead of district-by-year fixed effects. For the teacher-by-year fixed effects, we only use the student lagged outcome framework due to collinearity. While our test score results are attenuated from the inclusion of school-by-year fixed effects, there is still a significant increase in student test scores.

We find that the reduction in suspensions and absences is also robust. We see a similar pattern controlling for teacher-by-year-by-subject (classroom) fixed effects.

We expect these estimates to be attenuated because they explicitly restrict the mechanism through which a repeat teacher can "affect" her students. In particular, comparing students within classrooms would attenuate our estimates if non-repeat students either directly or indirectly benefit from having repeat students in their class. If repeat interactions improve student-teacher match quality, teachers may not only be more effective with these repeat students but may be able to re-allocate additional time and effort to their new students. Improved achievement and behavior among repeat student may also have positive spillovers to their classmates via peer effects. Below we test for and find evidence of such positive spillovers.

Finally, we control for the teacher's measured value added (VAM) from the previous year to directly address any concerns that more effective teachers are more likely to repeat. This reduces our sample by about one-third because we limit our analysis to teachers in tested grades and subjects in the previous year. We compare these results to those from our preferred specification estimated on the same sample. Controlling for teacher VAM results in no change in the estimated effect of repeat teachers (comparing column 6 to 8 and column 7 to 9.)

5.2 Classroom Selection

Because most repeat student-teacher assignments come from teachers moving across grades or teaching multiple grades, both teachers and students may have the opportunity to "choose to repeat with only well-matched students or teachers. For example, if there are two fifth-grade teachers but only one had taught fourth grade the year before, it is possible that only the students with whom the teacher was effective will select into her class during fifth grade. To account directly for this type of selection, we limit our sample to schools in which there is a single teacher for a tested grade and subject; in such schools, students cannot select a specific teacher

and teachers cannot select specific students. Columns 3 and 4 of Table 6 show that the results are remarkably consistent to the preferred results in columns 1 and 2, although they are substantially less precise because only 10% of our preferred sample attend a school with only one teacher in the subject. We find consistent positive effects on test scores and, although we lose substantial power for the student behavior outcome estimates, we still find negative point estimates for both absences and suspensions. Additionally, limiting the sample to schools with a single teacher per grade skews our observations both to smaller schools and to earlier grades.

Just as students may follow a well-matched teacher, teachers may choose to follow a particularly well-matched cohort of students into a higher grade. To abstract from this type of teacher selection, we limit our sample of teachers to only those that are not teaching in a new grade or subject. In other words, we restrict our estimates to teachers that could not have chosen to follow a particularly well-matched cohort of students into a new grade because that teacher already teaches in that grade.[9] Again, we find largely consistent results as shown in columns 5-8 of Table 6. We find a positive and significant effect on test scores and a negative effect on both absences and suspension. In contrast to limiting the sample to schools with a single teacher in each grade, this restriction skews the sample toward middle and high schools.

6 Student Heterogeneity

Different students benefit from different types of teachers and have different learning styles. It is possible that having a repeat teacher is more effective for certain types of students relative to others. Put differently, match quality may be more dynamic for certain types of student-teacher matches than for others. We have already shown reasonably consistent impacts of having a repeat teacher across different grades and subjects. In Table 8, we expand that analysis to examine the differential effects by

[9] An additional benefit of limiting the sample to teachers who consistently teach in the same set of grades and subjects is that we control for the possibility that teachers only move into grades where they believe they will be more effective.

student characteristics, such as academic performance, race, and gender. We interact the characteristic of interest with our indicator for a repeat teacher.

We present our results in Table 7 , showing the overall effect for the reference group (top row) and the difference in impact for each other group. We find that the positive achievement effects of having a repeat teacher are driven by students in the top half of the test score distribution (third and fourth quartile in the previous year).[10] However, while achievement effects are concentrated among higher-performing students, lower-performing students appear to benefit more in behavioral outcomes from repeating with a teacher. This is consistent with existing evidence that teachers contribute to students' long-term outcomes not only through their effects on academic skills but also through social-emotional skills not captured by test scores (Jackson, 2018; Kraft, 2019). It also provides suggestive evidence that the improvement in student achievement is not simply driven by increased class time.

A further source of heterogeneity in the effect of a repeat teacher could result from differential effects by gender and race. Recent work on implicit bias suggests that teachers may enter the classroom with substantial biases against students of color (Papageorge, Gershenson, & Kang, 2020; Starck et al., 2020); if building a relationship with individual students helps to ameliorate such biases, match quality may be more malleable over time for students of color than for white students. To explore this possibility, we interact a set of indicators for a student's race and gender with our repeat variable in Panel B of Table 7. White female students see the largest gains in achievement test scores from having a repeat teacher. The difference is particularly stark for female students of color who appear to have virtually no benefit to their test scores from repeat teachers. These differential effects are not driven by the fact that white female students are higher achieving on average as they are robust to both controlling directly for students' lagged test scores and for student FE. However, patterns are very different for the behavioral outcomes. The

[10]This is contrary to what would be expected if our results were driven by mean reversion.

reduction in suspension and particularly absences from having a repeat teacher is highest for male students of color. In fact, having a repeat teacher reduces absences by approximately 3% for boys of color. The additional effect of repeat teachers on male students of color is particularly notable because there is extensive evidence that they are more likely to be suspended and disciplined than their white peers and that those suspensions can lead to substantial negative outcomes including incarceration and lower earnings later in life (Cuellar & Markowitz, 2015; Costenbader & Markson, 1998; Mendez, 2003 ; Bacher-Hicks, Billings & Deming, 2019).

6.1 Spillover Effects

As described above, there is substantial scope for positive spillovers on non-repeat students. Teachers with a high number of repeat students may be able to better focus on building relationships with the non-repeat students and may require less time at the beginning of the year to establish classroom protocols. In other words, they may be able to both reallocate effort to non-repeat students and to improve their match quality with these students more rapidly. To test for potential spillover, we interact our repeat indicator with an indicator for whether 50% or more of a teacher's students in that subject year are also repeat students and present the results in Table 8.[11] We find a smaller (0.007 SD) but positive and significant effect of having a repeat teacher on students in classrooms with fewer than 50% repeat students. We find a larger effect of being in a high-repeat classroom for non-repeat students (0.011 SD) and an even larger effect (0.018 SD) for repeat students that are in high-repeat classrooms.

These results suggest that intentional policies to match students and teacher together across multiple classrooms such as looping may have even larger effects than those captured by our primary estimate. Both repeat students and their non-repeat peers perform better academically in majority repeater classrooms. The relationship between classroom composition and behavioral outcomes is less clear. Having a

[11] Alternative specifications ranging from indicators from 20% repeat students to 80% show positive spillover for all levels 40% and above

repeat teacher uniformly reduces absences and suspension but it doesn't consistently differ by the percentage of repeat students in the class.

7 Teacher Heterogeneity

Student-teacher match is a two-sided enterprise, suggesting that the characteristics of teachers might also matter. For example, repeat interactions may be particularly valuable for less effective or less experienced teachers if it takes them longer to build strong relationships with students. On the other hand, more effective or more experienced teachers may be better able to reallocate classroom time to maximize their knowledge of student styles.

To test the possibility that the effect of a repeat teacher varies by teacher experience, we interact the indicator for repeat with indicators for different bands of experience.[12] As seen in Table 9, repeat teachers improve test scores and decrease absences about equally across experience categories.

On the other hand, we find some evidence that the effect of repeating is larger for low-value-added teachers. To examine teacher heterogeneity by measured value added we interact our repeat indicator with indicators for the teacher having been in the second, third, and fourth quartile of performance in the previous year as well as controlling directly and linearly for that teacher's prior year VAM score. The results for this interaction specification are in columns 1 and 2 of Table 9. We find the largest impacts for teachers in the lowest value-added quartile. This analysis speaks to a common concern of parents about looping: that having the same "bad" teacher two years in a row could magnify the negative effect of that teacher. We find no evidence that the negative effects of a low measured value added teacher are amplified by a second year with that teacher. In fact, there is a positive effect of repeating with a teacher across all value added quartiles.

[12] Repeat teachers must have at least one year of experience by definition

8 Conclusion

Match quality matters for employees and firms as well as for teachers and schools. We study one specific type of match quality within schools, student-teacher matches, and document how match quality is a dynamic rather than fixed characteristic. We find substantial evidence that having a repeat teacher improves students' achievement and behavioral outcomes. Across grades 4 – 11, students perform better on standardized tests in ELA and math in their second (or more) year with a teacher. Repeat teachers also improve student behavioral outcomes, reducing suspension rates, absences, and truancy.

There is also significant heterogeneity in the effect of repeat teachers across student demographic and achievement groups. Students who are already higher achievers and white girls have larger tests score gains in years with repeat teachers compared to their lower achieving classmates. However, the opposite is true for behavioral outcomes. Students with lower test scores have a greater reduction in absences and the probability of being suspended than their higher achieving peers. The reduction in the probability of suspension is mainly driven by male students of color who are more likely to be suspended overall.

Differences in teacher value added and experience play a limited role in the impact of repeating. While more experienced and higher value added teachers improve test scores by more than their less experienced and lower value added peers, there is no evidence that repeating with a lower performing or less experienced teacher magnifies the negative effects of that teacher. In other words, repeating with a low performing or inexperienced teacher is better than having a different low performing or inexperienced teacher.

Documenting the potential gains from student-teacher match quality is important because policies to improve match within schools hold great potential as a near costless intervention. The type of repeated student-teacher interaction we

identify is generally an unintentional side effect of teachers moving between grades and subjects or teaching multiple grades and subjects at the same time. Despite the idiosyncratic nature of these repeat interactions, they still have a positive effect on student test scores and behaviors.

Our results are consistent with the literature on the importance of student-teacher relationships in building academic and socioemotional skills. While we are unable to directly measure the strength of the student-teacher relationship, we believe that our indicator for a repeat student-teacher match is a good, if imperfect, proxy for that relationship. Beyond their first year together, students and teachers have had more time to get to know each other's teaching and learning styles as well as to develop a stronger more trusting relationship. This can be particularly important in high school when the average teacher has over 100 students per year. A student who has had a teacher previously may have the opportunity to develop a closer relationship and, as such, a stronger connection to school. This is consistent with our results that truancy and the probability of being suspended are lower for students with repeat teachers in middle and high school when students also have non-repeat teachers in other subjects.

Although the benefits of unintentional repeat student-teacher matches are relatively small, our estimates likely understate the effect of more intentional policies of repeat pairings between teachers and students such as looping. With intentional loops, teachers can realize other benefits, such as adjusting the content of classes over two years in order to maximize learning. That students, both those with repeat teachers and those without, have higher test scores in classrooms in which more than half the students are repeats is further suggestive evidence of additional benefits from planned looping.

There is significant churn in the classes that teachers teach and a substantial likelihood that teachers could encounter the same student more than once, especially in middle and high school. Schools have to take into account a number of compli-

cations in terms of scheduling and assigning students to classes. Our results suggest there are benefits to leveraging teachers that move up grades, or teach across grades to create more repeat student-teacher matches that may benefit both students and teachers.

References

[1] Akerlof, George A. and Rachel E. Kranton. 2002. "Identity and Schooling: Some Lessons for the Economics of Education." *Journal of Economic Literature*, 40(4): 1167-1201.

[2] Bacher-Hicks, Andrew, Stephen B. Billings, and David J. Deming. The School to Prison Pipeline: Long-Run Impacts of School Suspensions on Adult Crime. No. w26257. *National Bureau of Economic Research*, 2019.

[3] Bond, L., Butler, H., Thomas, L., Carlin, J., Glover, S., Bowes, G., & Patton, G. (2007). Social and school connectedness in early secondary school as predictors of late teenage substance use, mental health, and academic outcomes. *Journal of Adolescent Health,* 40(4), 357-e9.

[4] Burke, D. L. (1997). Looping: Adding Time, Strengthening Relationships. *ERIC Digest.*

[5] Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, 104(9), 2593-2632.

[6] Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of human Resources,* 41(4), 778-820.

[7] Costenbader, V., & Markson, S. (1998). School suspension: A study with secondary school students. *Journal of School Psychology*, 36(1),5982.

[8] Cuellar, A. E., & Markowitz, S. (2015). School suspension and the school-to-prison pipeline. *International Review of Law and Economics*, 43, 98-106.

[9] Cunha, Flavio, and James J. Heckman. "Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation." *Journal of human resources* 43.4 (2008): 738-782.

[10] Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter?. *American Economic Review,* 95(2), 158-165.

[11] Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *Journal of personality and social psychology*, 92(6), 1087.

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

[13] Eisenhower, A. S., Baker, B. L., & Blacher, J. (2007). Early student–teacher relationships of children with and without intellectual disability: Contributions of behavioral, social, and self-regulatory competence. *Journal of School Psychology,* 45(4), 363-383.

[14] Franz, D. P., Thompson, N. L., Fuller, B., Hare, R. D., Miller, N. C., & Walker, J. (2010). Evaluating mathematics achievement of middle school students in a looping environment. *School Science and Mathematics,* 110(6), 298-308.

[15] George, P. S., & Lounsbury, J. H. (2000). *Making Big Schools Feel Small: Multiage Grouping, Looping, and Schools-within-a-School.*

[16] Gershenson, S. (2016). Linking teacher quality, student attendance, and student achievement. *Education Finance and Policy,* 11(2), 125-149.

[17] Gershenson, Seth, Hart, Cassandra M. D., Hyman, Joshua, Lindsay, Constance A.,& Papageorge, Nicholas. 2018. The Long-Run Impacts of Same-Race Teachers. *NBER Working Paper No. 25254.*

[18] Gershenson, S., Holt, S. B., & Papageorge, N. W. (2016). Who believes in me? The effect of student–teacher demographic match on teacher expectations. *Economics of Education Review,* 52, 209-224.

[19] Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of public economics,* 95(7-8), 798-812.

[20] Hamre, B. K., & Pianta, R. C. (2001). Early teacher–child relationships and the trajectory of children's school outcomes through eighth grade. *Child development,* 72(2), 625-638.

[21] Hamre, B. K., & Pianta, R. C. (2006). *Student-Teacher Relationships.*

[22] Heckman, J.J., Stixrud, J. and Urzua, S., 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics*, 24(3), pp.411-482.

[23] Heckman, J.J., 2006. "Skill formation and the economics of investing in disadvantaged children." *Science*, 312(5782), pp.1900-1902.

[24] Heckman, J. J. (2000). Policies to foster human capital. *Research in economics,* 54(1), 3-56.

[25] Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives,* 2(3), 172-177.

[26] Hill, A. J., & Jones, D. B. (2018). A teacher who knows me: The academic benefits of repeat student-teacher matches. *Economics of Education Review,* 64, 1-12.

[27] Cunha, F. and Heckman, J., 2007. "The Technology of Skill Formation." *The American Economic Review*, 97(2), p.31.

[28] Imberman, S. A. (2011). Achievement and behavior in charter schools: Drawing a more complete picture. The Review of Economics and Statistics, 93(2), 416-435.

[29] Jackson, C. K. (2009). Student demographics, teacher sorting, and teacher quality: Evidence from the end of school desegregation. *Journal of Labor Economics,* 27(2), 213-256.

[30] Jackson, C. K. (2013). Match quality, worker productivity, and worker mobility: Direct evidence from teachers. *Review of Economics and Statistics,* 95(4), 1096-1116.

[31] Jackson, C. K. (2014). Teacher quality at the high school level: The importance of accounting for tracks. *Journal of Labor Economics*, 32(4), 645-684.

[32] Jackson, C. K. (2018). What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes. *Journal of Political Economy*, 126(5), 2072-2107.

[33] Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of political economy*, 87(5, Part 1), 972-990.

[34] Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of school health,* 74(7), 262-273.

[35] Kraft, M. A., Blazar, D., & Hogan, D. (2016). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*

[36] Kraft, M.A. & Blazar, D. (2017). Individualized coaching to improve teacher practice across grades and subjects: New experimental evidence. *Educational Policy*, 31(7), 1033-1068.

[37] Kraft, Matthew A., and Sarah Grace. (2016) Teaching for tomorrow's economy? Teacher effects on complex cognitive skills and social-emotional competencies. *Brown University Working Paper.*

[38] Kraft, M.A. (2019). Teacher effects on complex cognitive skills and social-emotional competencies. *Journal of Human Resources*, 54(1), 1-36.

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

[40] Liu, J., & Loeb, S. (2019). Engaging teachers: Measuring the impact of teachers on student attendance in secondary school. *Journal of Human Resources*, 1216-8430R3.

[41] Loeb, S., Soland, J., & Fox, L. (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.

[42] Mendez, L. M. R. (2003). Predictors of suspension and negative school outcomes: A longitudinal investigation. *New Directions for Youth Development*, 99.

[43] Mincer, Jacob. 1974. *Schooling, Experience, and Earnings*. Columbia University Press.

[44] Murawski, W. W., & Lee Swanson, H. (2001). A meta-analysis of co-teaching research: Where are the data?. *Remedial and special education,* 22(5), 258-267.

[45] Ost, B. (2014). How do teachers improve? The relative importance of specific and general human capital. *American Economic Journal: Applied Economics,* 6(2), 127-51.

[46] Owens, Jayanti, and Sara S. McLanahan (2020). Unpacking the drivers of racial disparities in school suspension and expulsion. *Social Forces* 98, no. 4 : 1548-1577.

[47] Papageorge, N. W., Gershenson, S., & Kang, K. M. (2020). Teacher expectations matter. *Review of Economics and Statistics*, 102(2), 234-251.

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

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

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

[51] Starck, Jordan G., Travis Riddle, Stacey Sinclair, and Natasha Warikoo. "Teachers are people too: Examining the racial bias of teachers compared to other American adults." *Educational Researcher* (2020): 0013189X20912758.

[52] Todd, P. E., & Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113(485).

[53] Wedenoja, L. (2017). The Dynamics of High School Dropout. *Working Paper*

[54] Wedenoja, L. (2018). Beyond Averaged Yearly Attendance and Chronic Absenteeism: How Within-Year Attendance Patterns Predict High School Achievement. *Working Paper*

9 Tables

Table 1: Repeater and Non-Repeater Characteristics

Students	All Teachers	>10% Repeat	<10% repeat	Difference	
FRPL	0.488	0.501	0.488	0.013	***
SPED	0.069	0.068	0.069	-0.001	***
Female	0.499	0.499	0.499	0.000	***
LEP	0.030	0.018	0.031	-0.013	***
White	0.683	0.780	0.679	0.101	***
AfAm	0.231	0.161	0.234	-0.073	***
Hispanic	0.062	0.043	0.063	-0.020	***
Asian	0.018	0.010	0.018	-0.008	***
Lagged Math	0.070	0.036	0.072	-0.036	***
Lagged Reading	0.060	0.029	0.061	-0.032	***
Lagged Suspend	0.112	0.106	0.112	-0.006	***
Lagged Absent	7.690	7.440	7.700	-0.260	***
Lagged Truant	2.837	2.570	2.840	-0.270	***
Teachers	All Teachers	>10% Repeat	<10% repeat	Dfference	
Experience	11.35	12.31	11.22	1.086	***
White	0.864	0.883	0.862	0.021	***
Black	0.130	0.112	0.132	-0.020	***
Hispanic	0.002	0.001	0.002	-0.001	**
Age	41.57	42.40	41.44	0.956	***
Female	0.857	0.832	0.861	-0.029	***
Masters Degree	0.569	0.603	0.565	0.038	***
Lagged VAM	0.011	0.003	0.011	-0.008	***

Table contains all student-by-year observations in which a student is taking one of the following courses: Math 4 - Math 8, Reading 4 - Reading 8, Algebra II, Geometry, English II, English III. "Repearters" have a repeat teacher in at least one of those subjects, "Nonrepeaters" have no repeat teachers. The raw difference column is the the mean difference in characteristic between repeaters and non-repeaters with a two mean t-test. The w/in school difference column controls for school by year fixed effects. Significance levels: *0.1 **0.05 ***.01

Table 2: Repetition by Grade and Subject

TCAP Math Tests

	Students		Teachers			
Subject	N	% Repeat	N	Any%	10%+	50%+	90%+
4th Grade	578,068	0.023	27,302	0.060	0.054	0.020	0.002
5th Grade	572,874	0.034	21,718	0.094	0.083	0.035	0.007
6th Grade	570,421	0.027	12,677	0.092	0.076	0.045	0.011
7th Grade	563,728	0.043	11,571	0.160	0.141	0.086	0.022
8th Grade	527,144	0.110	11,041	0.328	0.287	0.177	0.053
All TCAP Math	2,812,235	0.046	84,309	0.119	0.104	0.055	0.013

TCAP Reading/ELA Tests

	Students		Teachers			
Subject	N	% Repeat	N	Any%	10%+	50%+	90%+
4th Grade	607,206	0.022	29,192	0.061	0.056	0.023	0.003
5th Grade	636,073	0.034	24,685	0.094	0.084	0.037	0.007
6th Grade	718,405	0.027	17,287	0.087	0.075	0.044	0.010
7th Grade	716,328	0.051	15,590	0.185	0.164	0.102	0.026
8th Grade	703,402	0.109	14,748	0.327	0.296	0.186	0.059
All TCAP Reading	3,381,414	0.049	101,502	0.128	0.115	0.063	0.016

High School EOC Tests

	Students		Teachers			
Subject	N	% Repeat	N	Any%	10%+	50%+	90%+
Algebra II	402,854	0.055	7,543	0.303	0.193	0.040	0.009
Geometry I	171,141	0.052	3,014	0.305	0.209	0.068	0.019
English II	719,320	0.056	13,595	0.243	0.188	0.075	0.023
English III	390,707	0.095	7,656	0.364	0.298	0.121	0.040
All EOC	1,684,022	0.064	31,808	0.3025	0.2229	0.074	0.016

Columns "N' and" "% Repeat" under student are the total number of student-year observations for each subject and the percentage of those students that have a repeat teacher in a different grade or EOC subject. Columns "N," "Any," "10%," "50%," and "90%" under teacher contain the number of teachers in each tested subject and the percentage of those teachers who have any, 10%+ , 50%+ , and 90%+ repeat students for that subject respectively.

Table 3: Paths of Repeat Teachers

Subject	Number of Teachers	Teachers Multiple Grades	New to Grade or EOC test	Same Tests Last Year	Stays in Grade Next Year	Stays in New Grade
			TCAP Math Tests			
4th Grade	874	15.3%	86.5%	9.5%	67.0%	66.2%
5th Grade	1,034	26.2%	66.8%	26.0%	67.3%	62.3%
6th Grade	522	52.5%	52.7%	37.9%	63.2%	53.8%
7th Grade	875	74.4%	37.8%	49.8%	75.3%	60.9%
8th Grade	1,708	71.7%	30.6%	59.5%	76.3%	62.1%
			TCAP Reading/ELA Tests			
4th Grade	963	19.1%	85.6%	10.3%	63.6%	63.1%
5th Grade	1,204	29.2%	66.9%	26.7%	62.7%	60.4%
6th Grade	704	58.5%	48.6%	41.8%	62.8%	53.9%
7th Grade	1,376	76.9%	38.7%	51.2%	68.7%	54.2%
8th Grade	2,369	72.4%	30.9%	60.7%	73.7%	60.5%
All TCAP	11,611	54.0%	49.9%	41.8%	69.5%	60.7%
			High School EOC Tests			
Algebra II	1,026	58.3%	43.8%	50.2%	75.0%	69.9%
Geometry	325	71.7%	72.6%	23.1%	70.5%	69.5%
English II	1,620	71.2%	36.0%	55.8%	71.1%	60.1%
English III	1,500	69.0%	42.3%	52.3%	72.7%	66.3%
All HS EOC	4,450	67.5%	42.5%	51.0%	72.5%	65.6%

Column 1 indicates the tested grade and subject for teachers with at least 10% repeat students with available data for the year before and after they are repeat teachers. Column 2 is the number of teacher-years in which the teacher has at least 10% repeat students. Column "Teaches multiple grades" is the percent of repeat teachers that have students in multiple grades (for grades 4-8) or tests (EOC subjects in high school.) "New to grade or test" indicates the percentage of teachers who were not teaching in this grade/subject during the prior year, "same tests last year" indicates that teachers taught the exact same tests in the previous year (for example taught both 6th and 7th grade math both years.) "Stays in grade next year" indicates teachers that continue teaching in the subject the following year and "stays in new grade next year" is the percentage of teachers that were new to their grade that continue in that grade the following year."

Table 4: Effects of Repeat Student-Teacher Matches

	All Grades		Grades 3 - 8		High School	
	FE	Lagged	FE	Lagged	FE	Lagged
Panel A: Achievement Outcomes						
Math and Reading/ELA	0.019 ***	0.024 ***	0.012 ***	0.023 ***	0.017 ***	0.026 ***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
N	8,039,335	7,179,281	5,675,085	5,213,443	2,295,120	1,965,836
Math	0.032 ***	0.022 ***	0.024 ***	0.018 ***	0.041 ***	0.026 ***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.005)	(0.004)
N	3,364,006	3,156,440	2,358,861	2,356,010	733,428	800,424
Reading/ELA	0.015 ***	0.025 ***	0.013 ***	0.026 ***	0.007 **	0.021 ***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)
N	4,300,688	3,963,516	2,878,113	2,811,876	1,252,531	1,151,634
Panel B: Behavioral Outcomes						
ln(absent+1)	-0.005 ***	-0.013 ***	-0.001	-0.013 ***	-0.006 ***	-0.015 ***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)
Ever Suspended	-0.010 ***	-0.017 ***	-0.006 ***	-0.017 ***	-0.003 **	-0.012 ***
	(0.001)	(0.001)	(0.001)	(0.00)	(0.001)	(0.001)
ln(truancies+1)	–	–	–	–	-0.003	-0.016 ***
					(0.002)	(0.004)
N	8,039,335	7,179,281	5,675,085	5,213,443	2,295,120	1,965,836
Lagged Outcomes		X		X		X
Student FE	X		X		X	
District x Year FE	X	X	X	X	X	X

Table contains main results for the effect of having a repeat teacher on students' cognitive and non-cognitive outcomes. Columns 1&2 contain pooled estimates for all grades, 3&4 for grades 3-8, and columns 5&6 for High School EOC exams along. Panel A includes cognitive results for high pooled math and reading, math alone, and reading/ELA alone. Panel B includes non-cognitive results for ln(1+total absences), ln(1+total truancies), and an indicator for whether the student had been suspended during the year. All specifications include controls for class and school level demographics including race, gender, ELL status, FRPL status, and lagged test scores, absences, and suspension as well as teacher experience controls. All columns also include district-by-year FE. Lagged columns also contain controls for student demographics, a cubic in lagged test scores (or 8th grade test scores for high schol students), and lagged absences and suspension. FE columns also contain student fixed effects and time-varying student controls. Standard errors in parentheses. Significance levels: *0.1 **0.05 ***.01

Table 5: Robustness Checks

	Preferred		Alternative FE			Teacher Value Added Sample			
	FE	Lagged	School by Year	School by Year	Teacher by Year	Preferred FE	Preferred Lagged	VAM Controls FE	VAM Controls Lagged
Reading and Math Scores	0.019 *** (0.001)	0.024 *** (0.001)	0.008 *** (0.001)	0.011 *** (0.002)	0.012 *** (0.002)	0.009 *** (0.001)	0.015 *** (0.001)	0.009 *** (0.002)	0.014 *** (0.002)
ln(absent+1)	-0.005 *** (0.00)1	-0.013 *** (0.001)	0.001 (0.001)	-0.024 *** (0.002)	-0.023 *** (0.002)	-0.009 *** (0.002)	-0.024 *** (0.002)	-0.009 *** (0.002)	-0.012 *** (0.003)
Ever Suspended	-0.010 *** (0.001)	-0.017 *** (0.001)	0.003 *** (0.001)	-0.007 *** (-0.004)	-0.007 *** (0.001)	0.002 *** (0.001)	-0.007 *** (0.001)	0.002 *** (0.001)	-0.007 *** (0.001)
N	8,039,335	7,179,281	8,039,335	7,179,281	7,179,281	5,714,747	”5,181,662 ”	5,714,747	”5,181,662 ”
Lagged Outcomes		X		X	X		X		X
Student FE	X		X			X		X	
District x Year FE	X					X	X	X	X
Teacher x Year FE					X				
School x Year FE			X	X					
VAM Controls								X	X

Columns 1&2 contain the preferred specifications from Table 4. Columns 3&4 are identical to 1&2 but control for school-by-year rather than district-by-year fixed effects. Column 5 controls for teacher-by-subject-by-year fixed effects in the lagged outcome specification. Columns 6-9 are specifications for teachers that have lagged value added scores. Columns 6&7 are the preferred specification on that sample and Columns 8&9 control directly for the teacher's prior year VAM estimate. Standard errors in parentheses. Significance levels: *0.1 **0.05 ***.01

Table 6: Selection

	Preferred		One Teacher per Grade		Same Grades Last Year		This Grade Last Year	
Panel A: All Grades Math and Reading/ELA Teachers								
	FE	Lagged	FE	Lagged	FE	Lagged	FE	Lagged
Scores	0.019 ***	0.024 ***	0.013 ***	0.020 ***	0.011 ***	0.016 ***	0.009 ***	0.017 ***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
ln(absent+1)	-0.005 ***	-0.013 ***	-0.004	-0.037 ***	-0.011 ***	-0.028 ***	-0.010 ***	-0.028 ***
	(0.001)	(0.001)	(0.004)	(0.005)	(0.003)	(0.003)	(0.002)	(0.003)
Ever Suspended	-0.010 ***	-0.017 ***	0.000	-0.009 ***	0.001	-0.009 ***	0.001	-0.008 ***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
N	8,039,335	7,179,281	973,197	986,890	5,300,457	4,498,884	5,343,669	4,876,173
Panel B: High School EOC Math and ELA Teachers								
ln(truant+1)	-0.003	-0.016 ***	0.004	-0.040 ***	0.000	-0.018 ***	-0.006	-0.021 ***
	(0.002)	(0.004)	(0.011)	(0.011)	(0.004)	(0.005)	(0.004)	(0.005)
N	2,295,120	1,965,836	114,991	142,176	1,247,208	1,178,090	1,380,553	1,277,974
Lagged Outcomes		X		X		X		X
Student FE	X		X		X		X	
District x Year FE	X	X	X	X	X	X	X	X

The preferred columns are the results from table 4. "One teacher per grade" contains the subsample of schools for which there is only a single teacher for each tested grade/subject. "Same tests last year" contains the subsample of teachers that taught the same combination of grades/subjects during the previous year. "This test last year" contains the subsample of teachers that taught in this specific subject in the previous year regardless of the other subjects they taught in that year. Standard errors in parentheses. Significance levels: *0.1 **0.05 ***.01

Table 7: Student Heterogeneity

	Test Scores		Absences		Ever Suspended	
	FE	Lagged	FE	Lagged	FE	Lagged
			Panel A: Student Test Score Quartile			
Repeat	-0.002	-0.003	-0.013 ***	-0.035 ***	0.004 ***	-0.017 ***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.001)	(0.001)
Repeat X	0.002	0.002	0.004	0.014 ***	-0.004 **	0.007 ***
Second Quartile	(0.003)	(0.003)	(0.003)	(0.003)	(0.001)	(0.002)
Repeat X	0.015 ***	0.016 ***	0.004	0.017 ***	-0.005 ***	0.012 ***
Third Quartile	(0.003)	(0.003)	(0.003)	(0.003)	(0.001)	(0.002)
Repeat X	0.022 ***	0.043 ***	-0.001	0.017 ***	-0.001	0.023 ***
Fourth Quartile	(0.003)	(0.003)	(0.003)	(0.004)	(0.001)	(0.002)
			Panel B: Race and Gender			
Repeat	0.017 ***	0.014 ***	0.002	-0.014 ***	0.002 **	0.002 **
	(0.002)	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)
Repeat X	-0.008 ***	-0.001	-0.009 ***	-0.006 **	0.002	-0.013 ***
White Male	(0.002)	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)
Repeat X	-0.017 ***	-0.015 ***	-0.014 ***	-0.015 ***	0.003 *	-0.014 ***
Female of Color	(0.003)	(0.003)	(0.004)	(0.004)	(0.002)	(0.002)
Repeat X	-0.011 ***	-0.001	-0.034 ***	-0.039 ***	-0.004 ***	-0.018 ***
Male of Color	(0.003)	(0.004)	(0.004)	(0.004)	(0.002)	(0.002)
Lagged Outcomes	X		X		X	
Student FE		X		X		X
District x Year FE	X	X	X	X	X	X

Each column is a regression that includes an indicator variable "Repeat" that equals 1 when a student is repeating with a teacher and interaction variables that equal 1 when a student both has a repeat teacher and is a member of the indicated group. The "Repeat X" effect is the marginal effect of repeating for a member of the indicated group, compared to the effect of repeating for the omitted group (first quartile students and white female students, respectively.) Standard errors in parentheses. Significance levels: *0.1 **0.05 ***.01

Table 8: Spillover Effects

	Test Scores		Absences		Ever Suspended	
	FE	Lagged	FE	Lagged	FE	Lagged
Repeat	0.007 ***	0.006 ***	-0.009 ***	-0.016 ***	0.002 **	-0.005 ***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
over 50% repeat	0.018 ***	0.011 ***	-0.006	0.006	-0.007 ***	-0.002
	(0.003)	(0.004)	(0.004)	(0.004)	(0.002)	(0.002)
over50%repeat	0.011 ***	0.018 ***	0.004	-0.027 ***	0.002	-0.006 ***
X repeat	(0.003)	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)
Lagged Outcomes	X		X		X	
Student FE		X		X		X
District x Year FE	X	X	X	X	X	X

Repeat is an indicator that a student has had a teacher in the past, over 50% repeat is an indicator that a student is in a classroom with more than 50% repeat students whether or not the student herself has had the teacher before, and over 50% repeat X repeat is an indicator that a student is in a classroom with over 50% repeaters and is a repeater herself. Standard errors in parentheses. Significance levels: *0.1 **0.05 ***.01

Table 9: Teacher Heterogeneity

Value Added Quartile Interaction				Teacher Experience Interaction					
	Test Scores			Test Scores		Absences		Suspensions	
	Lagged	FE		Lagged	FE	Lagged	FE	Lagged	FE
Repeat	0.030 ***	0.002	Repeat	0.010 ***	0.007 ***	-0.018***	-0.004 ***	-0.006 ***	0.004 ***
	(0.003)	(0.003)		(0.003)	(0.002)	(0.005)	(0.003)	(0.001)	(0.001)
Repeat X	-0.013 ***	0.004	Repeat X	0.003	0.003	-0.009 *	-0.006 *	-0.005 ***	-0.003 **
Q2 Teach	(0.003)	(0.003)	5-9 years	(0.003)	(0.003)	(0.005)	(0.003)	(0.002)	(0.001)
Repeat X	-0.016 ***	0.011 ***	Repeat X	0.003	0.005 *	-0.004	-0.005 *	0.003	0.001
Q3 Teach	(0.003)	(0.003)	10-17 years	(0.003)	(0.003)	(0.005)	(0.003)	(0.002)	(0.001)
Repeat X	-0.033 ***	0.024 ***	Repeat X	-0.001	0.004	-0.003	-0.003	-0.002	-0.003 **
Q4 Teach	(0.004)	(0.003)	18+ years	(0.003)	(0.003)	(0.005)	(0.003)	(0.002)	(0.001)
N	5,181,662	5,714,747		7,179,281	8,039,335	7,179,281	8,039,335	7,179,281	8,039,335
Lagged Out	X		Lagged Out	X		X		X	
Student FE		X	Student FE		X		X		X
Dist x Year FE	X	X	Dist x Year FE	X	X	X	X	X	X

Repeat is an indicator that a student is repeating with a teacher. Repeat X group are indicators that a student is repeating with a teacher and that teacher belongs to that group. The repeat X variables are the marginal, or additional effect of repeating with a teacher that is a member of the indicated group relative to the omitted group, teachers in the first quartile of measured VAM and teachers with less than 5 years of experience respectively. Standard errors in parentheses. Significance levels: *0.1 **0.05 ***.01