A teacher who knows me: The academic benefits of repeat student-teacher matches

Andrew J. Hill [a]
University of South Carolina

Daniel B. Jones [b]
University of South Carolina

Abstract. We provide new empirical evidence that increased student-teacher familiarity improves academic achievement. Drawing on rich administrative data from North Carolina elementary schools, we observe small but significant test score gains for students assigned to the same teacher for a second time in a higher grade. We control for selection into repeat student-teacher matches with both student and teacher fixed effects. The effects are largest for minority students, and there is no heterogeneity by teacher quality (as measured by value-added). We also provide suggestive evidence of spillover benefits: students assigned to classes in which a large share of classmates are in repeat student-teacher matches experience gains even if not previously assigned to that teacher themselves. This suggests that effects at least partly operate through improvements in the general classroom learning environment. Overall, our finding indicates that there may be potential low-cost gains from the policy of "looping" in which students and teachers progress through early school grades together.

[a] Hill: University of South Carolina, Darla Moore School of Business, Department of Economics, 1014 Greene St., Columbia, SC 29208. Email: andrew.hill@moore.sc.edu.
[b] Jones: University of South Carolina, Darla Moore School of Business, Department of Economics, 1014 Greene St., Columbia, SC 29208. Email: daniel.jones@moore.sc.edu.

1. Introduction

Among the many potentially important inputs into the education production function[1], few have proven to be as consistently important as teachers. For instance, both Rockoff (2004) and Rivkin et al. (2005) find large gains from being assigned to a higher quality teacher, much larger, in fact, than gains from reductions in class size. Chetty et al. (2014a, 2014b) show that being assigned to a higher quality teacher not only leads to higher test scores for students in grades 3 to 8, but also leads to better outcomes later in life. Understanding *how* teachers matter is clearly important, and the literature has explored several teacher-related factors relevant in education production (see, for example, Clotfelter et al, 2007). We propose a new mechanism that has not been directly explored, but is easily affected by policy: student-teacher familiarity from repeat student-teacher matches.

The potential importance of student-teacher interactions has been highlighted in previous research. Several papers show that students who share demographic characteristics with their teachers perform better (Dee (2005), Hoffman and Oreopoulos (2009), Fairlie et al (2014), Ouazad (2014), Gershenson et al (2016) and Egalite et al (2015)). Several recent studies document that students show higher levels of achievement when their teachers are relatively optimistic about their future outcomes (Papageorge et al. (2016), Hill and Jones (2017)). Collectively, these papers suggest that familiarity with or understanding of students may play an important role in determining teachers' impacts on students.

Our paper addresses this possibility directly. Specifically, we provide evidence on the academic gains resulting from more thoroughly established student-teacher relationships in grades 3 to 5, early school grades in which students are relatively young and teachers often provide emotional support in addition to instruction. We draw on administrative data on elementary school students and teachers in North Carolina. These data allow us to both observe the universe of teachers and students in the state for a number of years and match students to teachers. We identify students who – during the years we observe them – are assigned to the same teacher at least twice in different grades. We then estimate rich fixed effects models at the student-by-grade level (incorporating both student and teacher fixed

[1] See, for instance, Hanushek (2003) for an analysis of the efficacy of a variety of inputs in the education production function.

effects in our most complete models) to assess the impacts of being assigned to a teacher for the second time, taking performance on standardized end-of-grade tests as our outcome. The inclusion of student fixed effects means we are identifying within-student growth in achievement in the second year that a student is assigned to the same teacher. We anticipate higher student test scores in the second year of the match compared to the first year if the establishment of student-teacher relationships impacts student achievement.

Ultimately, we find clear evidence that students perform better in the second year they are matched to a particular teacher. The estimated gain from a repeat student-teacher match ranges from 0.02 to 0.12 test score standard deviations. The lower end of that range stems from specifications with both student and teacher fixed effects, a rich model that removes a considerable amount of variation. Still, for the sake of comparison, this effect is similar in magnitude to the conditional black-white achievement gap in the same North Carolina administrative data[2]. Importantly, we rule out that results are driven by selection of students into repeat matches with teachers with whom they performed well in in the past.

We present a series of results documenting heterogeneity in our main result. First, we find that the benefits of being rematched with a teacher are not dependent on teacher quality (as measured by teacher value-added). And, second, we test how our result varies for different ethnic groups, finding that results are strongest for minority students. We think this is unsurprising given the average teacher is not from a minority group and therefore the initial "social distance" between a minority student and teacher may be relatively high, meaning there is more to gain from increased student-teacher familiarity.

While we study repeat student-teacher matches primarily as a window into the importance of student-teacher relationships, our results shed light on a policy which has received relatively little attention in the economics of education literature: "looping". Looping, a policy in which whole classes (or most of the students within a class) are taught by the same teacher at least two years in a row, has been adopted by education policymakers throughout the world for decades (Cistone & Shneyderman, 2004). Cistone and Shneyderman (2004) conduct a survey of teachers in schools with looping and find that a large majority of teachers agree with the statement "Looping enhances the quality of the working relationships between teachers and students."

---

[2] Hill & Jones (2016) estimate the size of the black-white achievement gap conditional on a variety of available controls, including lagged test scores, gender, race, neighborhood poverty status, and parental education.

Our results suggest that, at least with respect to gains in tests scores, looping is a worthwhile policy. While there is no official looping policy in our data, in several cases sizable fractions of a class in some year *t* are rematched with the same teacher in year *t+1*. We find that in such settings, the "unofficial" looping we observe has clear benefits not just for the students who were rematched, but also for other students in the class; repeat student-teacher matches appear to generate spillover benefits to other students in the class. Although our results on this are more suggestive than conclusive, this indicates that repeat student-teacher matches may improve the general classroom learning environment rather than generate gains solely through more productive student-specific interactions with the teacher for rematched students.

2. Empirical methodology

We estimate the effect of repeat student-teacher matches on academic achievement using the following model:

$$Y_{ijgt} = \theta RepMatch_{ijt} + \alpha_i + \beta_j + \delta_g + \eta_t + \varepsilon_{ijgt}$$

Y_{ijgt} is the end-of-grade test score of student i taught by teacher j in grade g during year t. It is standard normalized so that the magnitude of the estimated treatment effect is in units of standard deviations. The treatment variable $RepMatch_{ijt}$ indicates whether student i was also taught by teacher j in a previous grade and year (before grade g and year t).

A student who is repeating a grade with the same teacher is not considered a repeat match. We do not want to consider these students as treated in our analysis because grade repetition has its own effect on school performance,[3] and we do not want to conflate our estimates of the effect of repeat student-teacher matches with this effect. Given that our analysis focuses on students in grades 3 to 5, $RepMatch_{ijt}$ can therefore only be one for 4th and 5th graders.

Students and teachers are not experimentally assigned to repeat student-teacher matches. It is therefore critical to control for factors that may affect the probability of

[3] See, for example, Eide and Showalter (2001), Jacob and Lefgren (2009), Manacorda (2012), and Hill (2015).

students and teachers being involved in a repeat match in order to identify a causal effect of repeat matches. There are two first-order concerns.

First, the ability of students may be correlated with the likelihood of being involved in a repeat student-teacher match. For example, teachers who already teach different grades in different years may ensure that they are disproportionately assigned the better students when they teach a higher grade (who they can easily identify given they have recently taught them), or, alternatively, teachers may be more likely to teach a class in the subsequent grade when they have an above average class in the current grade. This would upwardly bias estimates of the effect of repeat student-teacher matches on school performance. On the other hand, administrators may try to reduce the burden of difficult or less able students by disproportionately assigning them to the same teacher in multiple years if there is a belief that difficult students are more easily managed when they are known by the teacher.

We control for potential confounders related to student ability by including student fixed effects α_i in the model, which also deal with any other observable or unobservable student characteristics that may be correlated with the likelihood of being involved in a student-teacher repeat match.

Second, teacher quality may be correlated with the probability of being in a repeat student-teacher match. For example, parents may be more likely to request the same teacher in multiple years when that teacher is of a high quality (and similarly more likely to request a new teacher when the previous teacher was of a low quality). Under this scenario, the estimated effect of repeat matches on academic achievement would also be biased upward. A related bias may arise from teacher sorting. Higher quality teachers may disproportionately sort into schools in which it is more common for teachers to teach multiple grades if teachers view this as an opportunity to build longer-term relationships with students and these teachers are, on average, of a higher quality than other teachers. This would also upwardly bias the estimated effect. Teacher fixed effects β_j are included in the model to control for any selection into repeat student-teacher matches related to teacher quality or other teacher characteristics.

Loosely speaking, the inclusion of student and teacher fixed effects means that the effect of repeat student-teacher matches – the coefficient on the $RepMatch_{ijt}$ indicator – is identified by comparing the end-of-grade test scores of the same student with the same teacher in a grade in which they have the same teacher for a second time to the grade in which they had the teacher for the first time.

Grade fixed effects δ_g and year fixed effects η_t control for any systematic variation in end-of-grade test scores across grades and time, as well any trends (over both time and grades) in the extent of looping observed in the data and academic achievement.

It is worth noting for expositional purposes that it would be possible to include student-by-teacher fixed effects in the model – a dummy variable for every student-teacher combination – but, as would be expected, this model turns out to be over-fitted and soaks up most of the variation in the data. The advantage of excluding these fixed effects, however, is that the less-specified model allows us to directly probe the extent of potential selection into repeat student-teacher matches in a way that would not be possible with student-by-teacher fixed effects.

We do so in some specifications by including an indicator $EverRepMatch_{ijt}$ in the model that is equal to one when students and teachers are in the first *and* repeating year of a repeat student-teacher match relationship. In other words, it is always one when $RepMatch_{ijt}$ is one, but is also one in the first year that student i is taught by teacher j. This variable could not be included in a model with student-by-teacher fixed effects as they would be perfectly collinear.

If the coefficient on $EverRepMatch_{ijt}$ is positive, it means that higher-performing students are more likely to be involved in repeat student-teacher matches, which would be evidence of positive selection and a potentially upwardly biased estimate of the treatment effect. We show in the results section how the coefficient on this variable changes with the inclusion of teacher and student fixed effects to better understand the nature of potential selection into repeat student-teacher matches, and, most importantly, whether our controls adequately deal with any selection into repeat matches.[4]

In addition to the primary sources of bias related to student ability and teacher quality discussed above, an additional factor that may be correlated with student

[4] It is worth noting that even if this coefficient were non-zero, it would still be possible to make the claim that the coefficient on the primary variable of interest $RepMatch_{ijt}$ identifies a causal effect of repeat student-teacher matches on school performance. In this case, the parameter of interest would be an estimate of the average achievement boost caused by repeat matches *for the typical positively or negatively selected student* rather than an average treatment effect. In order to generalize our results, though, the student and teacher fixed effects in the model would need to capture all the selection into repeat student-teacher matches.

achievement and the likelihood of being involved in a repeat student-teacher match is school size. For example, it may be the case that students in smaller schools are more likely to be matched to the same teacher in multiple grades if teachers in smaller schools either teach more grades or have multi-grade classrooms. This could be particularly true in rural schools. If school size is also correlated with academic achievement (see Kuziemko (2006) and Gershenson and Langbein (2015), for example), our estimated effect of repeat student-teacher matches may be biased. School fixed effects will not fully capture school size effects given schools grow and shrink over time, especially considering our study sample of almost two decades. In addition to showing results are robust to school fixed effects, we therefore test whether results are sensitive to the inclusion of two additional time-varying fixed effects controls: the number of teachers who teach the same grade as teacher j in the school that year, and class size. The first of these is likely highly correlated with school size, but is in many ways a more direct control for potential bias caused by having fewer potential teachers with whom to match in smaller schools, while the second is included to control for unusually small classes that may be indicative of particularly small schools or multi-grade classrooms. These specifications allow us to test whether any effect of repeat student-teacher matches persists not only when we are comparing the same student taught by the same teacher across two grades, but also a student who had the same number of potential other teachers in the school that year, and is in a class with the same number of classroom peers.[5]

3. Data

We use data on the population of 3rd to 5th grade students and their teachers in public North Carolina elementary schools between 1997 and 2013 obtained from the North Carolina Education Research Data Center (NCERDC). Students are matched to the teacher who administered the end-of-grade test, which, for the early grades studied in this paper, is almost certainly the teacher who taught the class. We focus on end-of-grade mathematics test scores in this paper, showing in Appendix Table 1 that the effects of repeat student-teacher matches are similar for reading.

[5] It is worth noting that we have estimated specifications where we restrict the sample to students more likely to eventually be engaged in repeat student-teacher match (based on a variety of observable characteristics of the student's school including urban-rural status of the school, number of students in the schools, etc.). Results are not reported but are very similar to our main results.

Descriptive statistics are reported in Table 1. We have 2,111,082 unique 3rd to 5th grade students and 70,471 unique teachers who teach at least one 3rd to 5th grade class during the study period. As expected, about half the students are female. Forty percent of the students are minorities, the majority of whom are black. Turning to teachers, more than ninety percent are female, which is unsurprising given we are considering elementary schools. Notably, the share of teachers who are minorities is considerably smaller than the share of students who are minorities. This observation is particularly relevant when we relate our findings to the "teacher like me" literature; there is clearly a shortage of minority teachers if there is a desire to match students to teachers of the same race. On average, teachers have nine years of experience in the data.

Only three percent of students experience a repeat student-teacher match in the 4th or 5th grade. (Recall that students matched to the same teacher while repeating a grade are not considered to be in a repeat match, so 3rd grade students can never be in a repeat match.) Although a small share, this represents over sixty thousand students given the large sample. Almost ten percent of teachers in the sample are ever involved in a repeat student-teacher match. This is considerably higher than that for students given that typical students are exposed to only three teachers in over the three-year period from 3rd to 5th grade, while teachers are exposed to about 72 students during the same three-year period given an average class size of 24 students.

The final panel on the right describes the more than 5 million student-grade observations. On average, the typical student has 4.7 potential teachers in a given grade at their school. The final four variables are indicators for "looping classes" – classes in which at least the given share of students in the class are in a repeat student-teacher match. We are interested in the extent to which repeat student-teacher matches are a product of traditional looping, the practice of entire classes of students progressing with their teachers to higher grades. In these types of classes, almost all the students in the class would be in repeat student-teacher matches. The descriptive statistics in Table 1 indicate that these classes are uncommon, and not the primary source of repeat student-teacher matches. Only 0.3 percent of student-grade observations are in classes with a looping share of eighty percent, while 2 percent are in classes with a looping share of at least twenty percent.

4. Results

4.1 Main results

Our main results are reported in Table 2. Column 1 reports results from the model without controls, showing that 4th and 5th graders in repeat student-teacher matches score 0.123 standard deviations higher than other students on end-of-grade tests, on average. School fixed effects are included in Column 2 to control for the possibility that schools with teachers consistently teaching different grades in different years – perhaps schools with more administrative flexibility – are systematically better than other schools. We still observe the positive relationship between school performance and repeat student-teacher matches, indicating that school-specific factors do not drive the association between repeat matches and achievement.

Columns 3, 4 and 5 include teacher fixed effects, student fixed effects, and both teacher and student fixed effects, respectively. As discussed in the empirical methodology section, we consider these to be the key controls for identifying the causal effect of repeat student-teacher matches on school performance.

Teacher fixed effects do not affect the magnitude of the estimated effect (0.129 in Column 3 in comparison to 0.123 in Column 1), suggesting that higher quality teachers are not disproportionately involved in repeat student-teacher matches. The inclusion of student fixed effects in Column 4, however, considerably attenuates the estimated effect. The coefficient of interest falls by about three quarters to 0.036. The considerable decrease in the estimate suggests that students are positively selected into repeat student-teacher matches. Strikingly, though, after controlling for selection into repeat matches on student ability, students in repeat student-teacher matches still significantly outperform other students.

Our preferred specification in Column 5 includes both student and teacher fixed effects. Loosely speaking, we can interpret the results from this column as saying that the same student taught by the same teacher scores 0.024 standard deviations higher on end-of-grade tests when they are in a repeat student-teacher match. We consider this to be a small, but nontrivial, effect. It is of a similar order of magnitude to the within-school conditional black-white achievement gap (see Hill and Jones (2016), for example), and is subsequently shown to be considerably larger than the "teacher like me" effect, which has received considerable attention in the literature both in terms of estimating the direct effect of having a teacher of the same race on test scores and understanding the mechanisms through which the effect operates.

School fixed effects are included in the model in Column 6 for completeness. These are identified from the relatively infrequent circumstances in which students and teachers move schools. The estimated parameter of interest is statistically unchanged.

In Appendix Table 2, we show that the effect is driven by rematches in sequential grades. This table reports results from two models: one specification in which matches are only defined if they occur in sequential grades (Panel A), and another specification that includes an interaction between the original repeat match indicator and an indicator for being in a sequential grade (Panel B). This finding suggests that some of the benefits from increased student-teacher familiarity may arise from having better information about a student's past academic achievement, although it is difficult to disentangle this channel from simply having better information about the student on non-academic dimensions.

We now probe whether our empirical specification reliably identifies a causal relationship between repeat student-teacher matches and achievement. In particular, Table 3 explores the extent to which student and teacher fixed effects deal with potential selection into repeat student-teacher matches. As discussed in the empirical methodology section, the coefficient on an additional variable $EverRepMatch_{ijt}$ – an indicator that is one for the first *and* repeating year of a student-teacher match – should be equal to zero in order for us to claim that assignment to repeat student-teacher matches is conditionally quasi-random.

The results in Column 1 of Table 3 include this additional variable, but exclude teacher and student fixed effects. This simple specification suggests positive selection into repeat matching. Students involved in repeat matches score 0.108 standard deviations higher on end-of-grade tests, on average, in both the first *and* repeating year of a match than other students. Interestingly, though, these students still experience an additional boost to achievement in the second year of a match of 0.017 standard deviations.

The subsequent columns of Table 3 investigate the extent to which the key fixed effects controls attenuate the coefficient on this "ever repeat match" indicator, which we interpret as a measure of how much selection into repeat student-teacher matches our richer models capture.

The inclusion of teacher fixed effects in Column 2 does not appear to affect the extent of selection into repeat matches; the coefficient on $EverRepMatch_{ijt}$ is unchanged. This suggests that the students ever involved in repeat matches with a given teacher are higher achieving, on average, than the students never involved in repeat matches with that teacher. We can hypothesize from this finding that teachers may have some ability to affect class assignment, ensuring that better students are assigned to their classes when they teach different grades in different years, or, alternatively, that they only teach multiple grades when they know they will have an above average class.

Student fixed effects are included in Column 3. The coefficient on the additional indicator falls considerably from 0.108 to 0.017. This tells us that estimating effects using within-student rather than across-student variation deals with a considerable share of the selection into repeat matches, which is consistent with student fixed effects having a considerable effect on the parameter of interest in Table 2. The coefficient remains positive and statistically significant, though, indicating that not all selection into repeat matches is based on student ability or other student characteristics. Although the remaining selection into repeat matching is relatively small, we learn from this that for any given student, the teachers ever involved in repeat matches are of a higher quality (or, alternatively, are more likely to be associated with some other positive shock to school performance), on average, than the teachers never involved in repeat matches with that student. One interpretation of this is that parents may have some ability to ensure that their children only repeat match with better teachers.

We include both student and teacher fixed effects in Column 4. The coefficient on the "selection" variable – the indicator for ever being in a repeat match with the current teacher – is now zero. Repeat student-teacher matching boosts achievement by 0.022 standard deviations in this model, which is very similar to the corresponding estimate of 0.024 from Table 2. Our takeaway from this specification is that student and teacher fixed effects adequately control for any selection into repeat student-teacher matches, allowing us to interpret estimates as causal average treatment effects.

Columns 5 and 6 of Table 3 show that results are robust to controlling for the number of teachers in the grade in the school, and class size. The first of these additional controls is meant to capture the possibility that students may both be more likely to be involved in repeat student-teacher matches and perform better in smaller schools with fewer teachers in every grade, while class size fixed effects ensure that the estimated effect is not driven by particularly small classes that may be indicative of rural schools or multi-grade classrooms. The estimated causal relationship between repeat student-teacher matches and the academic achievement of 4th and 5th graders is generally unaffected by these controls.

4.2 Heterogeneity by teacher quality

We now address a major concern related to repeat-student matches: what happens to students with low quality teachers who are then faced with low quality teachers for multiple years? Given the evidence that 3rd grade teachers have long-run effects on labor market outcomes (Chetty et al., 2011), parents may be justifiably concerned about repeat student-teacher matches having a particularly negative effect on student achievement when their children have low quality teachers. On the other hand, it may be the case the low quality teachers perform better when they know their students and have developed relationships with them. In this case, looping and repeat student-teacher matches may actually be a relatively low-cost tool to improve teacher performance for low quality teachers.

We probe heterogeneity in the effect of student-teacher matches by interacting the treatment variable with proxies for teacher quality. In particular, we construct a simple measure of time invariant teacher value-added (VAMs) by regressing student end-of-grade test scores on teacher fixed effects (and a variety of controls) using the sample of students not involved in repeat student-teacher matches.[6] We partition teachers into quartiles based on these teacher fixed effects, generating four teacher quality bins, and then fully interact our repeat-match indicator with the four teacher quality indicators. Note that it is important to control for the quality quartile of the student's teacher in the *previous* grade to ensure that we are not capturing the composite effects of previous teacher quality and current repeat student-teacher matching.

Results from modified specifications, where we interact our repeat-match indicator with teacher quality indicators, are reported in Table 4. Column 1 of Table 4 shows that the general estimated effect of repeat matches is smaller, but still positive, in the reduced sample of 4th and 5th grade students for whom teacher VAMs could be estimated for both their current and previous teachers. Turning to the interaction specification in Column 2, we find no evidence of heterogeneity in the effect of student-teacher repeat matches by teacher quality, although we do observe the expected achievement gains for students exposed to

---

[6] There is a large literature exploring the validity of teacher value added measures (Rothstein, 2010; Chetty et al, 2014). Given we are using our simple estimates as a coarse measure of teacher quality in this paper, we do not focus on these concerns here. As an aside, it is worth noting that the phenomenon of looping may generate its own concerns for generating valid measures of teacher value added (such as those for tracking (Jackson, 2014), for example).

higher quality teachers in the previous grade. Although low quality teachers are associated with lower academic achievement, the students of low quality teachers benefit just as much from repeat student-teacher matches as students of high quality teachers relative to what they would have achieved if they were assigned to a new teacher of the same quality. Interestingly, the estimated gain from being in a repeat match (0.021 standard deviations) is similar in magnitude to the effect of replacing a student's teacher from the previous grade who was in the lowest quality quartile with a teacher who was in the second highest (0.019) or highest (0.024) quality quartile.

We also explore differential effects of repeat student-teacher matches along dimensions of teacher experience and teacher performance on credential exams. A teacher's z-score is a normalized measure of his or her performance on teacher licensing exams. Results in Column 3 of Table 4 show that the effects of repeat student-teacher matches are not affected by this measure of teacher ability, while the estimates in Column 4 show that teacher experience similarly has no influence on the effect of looping.

4.3 Heterogeneity by student race and student-teacher race congruence

Our analysis of the effects of repeat student-teacher matches on academic achievement has considerable parallels with the extensive "teacher like me" literature.[7] This literature generally explores whether students perform better when they are exposed to teachers of the same race or gender. One of the channels through which student-teacher matches on observable characteristics may improve students' school performance is that teachers of the same race and gender may better understand the corresponding subset of students of that race or gender. Repeat student-teacher matches may allow teachers to develop a similar, if not better, understanding of their students than that provided by a race or gender match.

In addition, it may be the case that repeat student-teacher matches particularly benefit minority students. For example, deeper student-teacher relationships may make minority students feel more understood and included in the classroom. Furthermore, if black students are more likely to come from more challenging family environments (such as single

[7] See, for example, Dee (2005), Hoffman and Oreopoulos (2009), Fairlie et al (2014), Ouazad (2014), Gershenson et al (2016) and Egalite et al (2015).

parent households), then teachers with better understandings of their specific backgrounds may be able to more adequately or appropriately address needs arising outside the classroom.

Comparing and contrasting the magnitudes of potential achievement gains from repeat student-teacher matching and student-teacher race matching may also be interpreted through a policy lens. It is arguably much less costly to generate repeat student-teacher matches than student-teacher race matches. The prior requires teachers be certified to teach multiple grades – which is likely not too burdensome for the 3rd to 5th grade teachers studied in this paper, although, admittedly, may be more costly in terms of teacher training for higher grades – while the latter requires the composition of teachers to match the composition of students within schools – which may take both years of teacher training and targeted teacher sorting across schools. In addition, we question the desirability of promoting student-teacher race matches in the long run given the clear potential for widening educational inequality and other adverse consequences.

In this subsection, first, we investigate the potential for repeat matches to have differential effects on minority students who may have more to gain from improved student-teacher relationships. And, second, we compare and contrast the repeat student-teacher matching effect with the "teacher like me" effect.

Defining a student race indicator $Minority_i$, and the indicator $RaceMatch_{ijt}$ to be one when student i and teacher j are of the same race, our final specification investigates whether repeat student-teacher matches have different effects on minority and non-minority students, as well as compares and contrasts our findings with those from the related "teacher like me" literature.

$$\begin{aligned} Y_{ijgt} \quad = \quad & \theta RepMatch_{ijt} + \pi(Minority_i \times RepMatch_{ijt}) + \rho RaceMatch_{ijt} \\ & + \sigma(Minority_i \times RaceMatch_{ijt}) + \alpha_i + \beta_j + \delta_g + \eta_t + \varepsilon_{ijgt} \end{aligned}$$

For white students, the effect of repeat student-teacher matches on end-of-grade test scores is θ, while the effect of student-teacher race matches is ρ. For minority students, the corresponding effects are $(\theta + \pi)$ and $(\rho + \sigma)$. We report estimates for white students θ and minority students $(\theta + \pi)$ rather than estimates of differential effects π; the corresponding differential effects are reported in Appendix Table 3.

The estimates in Column 1 of Table 5 show that the effects of repeat student-teacher matches on end-of-grade test scores are positive for non-minority (white) and minority students, but, strikingly, the gains are significantly larger for minority students. We focus on the black-white difference in remaining columns so that estimates can be directly related to the literature on the black-white achievement gap, although the findings are very similar if we do not make this restriction.[8] In Column 2, we see that black students benefit more from repeat matches than white students, which just indicates that race differences in Column 1 are not driven by Hispanic students.

The model in Column 3 allows us to compare and contrast the effects of repeat student-teacher matches and the effects of same-race student-teacher matches. The boost to achievement from a repeat match is 0.028 standard deviations, while it is 0.007 standard deviations for a student-teacher race match (which is very similar to the corresponding student-teacher race match estimate in Egalite et al (2015) using similar Florida administrative data). Both are statistically significant, but repeat matches are shown to be more beneficial. Interestingly, there is no additional gain from experiencing both a repeat student-teacher match and a same-race student-teacher match.

The results in Column 4 explore differences in these effects for white and black students. There are a few noteworthy observations: first, the repeat match effect is only present for white students when there is also a same-race match (white students gain 0.034 standard deviations on end-of-grade test scores in this circumstance), although this should be interpreted with the consideration that the majority of teachers in the sample are white; second, the repeat match effect is larger than the "teacher like me" effect for both white and black students; and, third, the "teacher like me" effect is larger for black students than white students (0.012 in comparison to 0.006). The findings in Columns 1 to 4, first, show that repeat student-teacher matching is particularly beneficial to minority students, and, second, emphasize that the magnitude of the effect, although small, is nontrivial relative to other well-known effects associated with student-teacher matching.

4.4 "Looping" and probing the potential for spillovers resulting from repeat matches

[8] Restricting the sample to white and black students reduces the sample by about ten percent in North Carolina.

Repeat student-teacher matches may improve school performance because teachers get to know their students and are able to adjust and target their teaching styles appropriately. Similarly, students may adapt their learning approach to the teaching style of the teacher, although this is arguably less likely to be the case for the students in elementary grades studied in this paper. Less directly, teachers' better understanding of their students may also help them provide discipline and improve the classroom atmosphere. It is therefore plausible that repeat student-teacher matches not only improve outcomes for the students involved in repeat matches, but also other members of the class.

Spillovers of this form are only distinguishable if not all students in a given class are in a repeat student-teacher match. If repeat matches are typically generated by the traditional practice of looping – entire classes progressing to the next grade with the teacher –the opportunity to estimate spillovers will be minimal. Perhaps surprisingly, the majority of repeat student-teacher matches in the data do not occur for entire classes; in fact, only five percent of repeat student-teacher matches occur in classes where all other classmates are also looping.

We probe the potential for spillovers from repeat student-teacher matches in Table 6 by constructing an indicator for "looping classes" to include in the model. This variable is equal to one if the share of students who are involved in a repeat student-teacher match in the class exceeds some threshold share L. To be clear, this indicator will be switched on for any student who is currently in a class with at least the threshold share L of repeat matchers irrespective of whether the given student is in a repeat match themselves or not. Defining this variable as $LoopingClassL_{ijt}$, the below specification captures both the direct effect of being involved in a repeat student-teacher match and any potential spillovers from being in class with repeat matchers. We are now relying on student and teacher fixed effects to capture selection into both a repeat student-teacher match and a looping class.

$$Y_{ijgt} = \theta RepMatch_{ijt} + \mu_L LoopingClassL_{ijt} + \alpha_i + \beta_j + \delta_g + \eta_t + \varepsilon_{ijgt}$$

If both θ and μ_L are positive, there is evidence not only that students involved in repeat student-teacher matches show higher academic achievement, but also that all students in the class receive a performance boost when a sufficient share of their classmates are in repeat student-teacher matches.

Panel A of Table 6 reports achievement differences between students in looping classes and other classes. Column 1 indicates that students in classes in which at least 20 percent of their classmates are experiencing the teacher for the second time outperform students in other classes by 0.015 standard deviations, on average. Columns 2 to 4 show that this performance boost increases with the share of looping students in the class (40, 60 and 80 percent, respectively).

We cannot claim that this provides evidence of spillovers, of course, because the students actually in the repeat matches may experience all the gains. The model in Panel B – which also includes the original student-teacher repeat match indicator – allows there to be both a repeat match effect and a looping class effect. From Column 1, students in repeat student-teacher matches in classes in which at least 20 percent of their classmates are looping outperform other students by 0.021 standard deviations, but there is no indirect benefit from being in a class with other repeat matchers (an imprecisely estimated effect of 0.004 standard deviations). However, once the share of repeat matchers in the class reaches 40 percent (Column 2), there is both a direct gain from being in a repeat match and a spillover effect from being in a class with classmates who are in repeat matches. This gain persists when the threshold share for defining a looping class increases to 60 percent and 80 percent (Columns 3 and 4). The relative flatness of this estimate beyond the threshold share of 40 percent indicates that spillover effects do not appear to increase with the share of repeat matchers in the class once they are initiated.

Finally, Panel C of Table 6 allows potential spillover effects to differentially affect students who are in repeat matches and students who are not in repeat matches. Results reported in the first two columns of this panel suggest that achievement gains from repeat student-teacher matches only arise in classes with a sufficient share (20 or 40 percent) of other repeat matchers. There are no spillover effects for non-repeat matchers in these classes. Once 60 percent of students in a class are in repeat matches, though, Column 3 tells us that there is a direct benefit for repeat matchers (0.015 standard deviations), an indirect benefit for non-repeat matchers students (0.016 standard deviations), but no additional spillover effect if already in a repeat match (0.003 standard deviations). The results for those not in repeat matches are less clear when the share of repeat matchers reaches 80 percent of the class (Column 4), although there are now only a small number of non-repeat matchers in the class driving effects.

Overall, results in Table 6 are interpreted as suggestive evidence that repeat student-teacher matches generate benefits not only through specific student-teacher interactions, but also through changes in the general classroom learning environment, such as improvements in the classroom atmosphere or the enforcement of discipline. Spillover effects for non-repeat matchers appear to initiate when the majority of students in the class are in a repeat student-teacher match.

5. Conclusion

In this paper, we assess the impact of repeat student-teacher matches in grades 3 to 5 on academic achievement. Drawing on administrative data from North Carolina, we estimate rich fixed effects models and find that students who are matched to a particular teacher for a second time score higher on standardized end-of-grade tests than they did in their first pairing with the same teacher. Student-teacher familiarity improves student achievement.

These results shed light on the importance of student-teacher relationships in determining academic performance. We use repeat student-teacher matches as a window into the importance of teacher familiarity with students, but there are, of course, many other ways that teachers may have more established relationships or greater familiarity with certain students. For instance, smaller class sizes may allow teachers to more fully invest in individual students. Similarly, as discussed above, our results are linked to the literature on student-teacher demographic congruence (a "teacher like me"). This literature shows that students who share demographic characteristics with their teachers perform better. While we certainly do not claim that our results can fully explain either the benefits of smaller classes or having a "teacher like me", the clear benefit of student-teacher familiarity we document in this paper may partly explain the mechanism through which these treatments impact student achievement.

Finally, our results also speak to a policy which has received little attention in the economics of education literature: "looping". As noted in the introduction, looping is the practice of assigning entire classes (or most of the students from a class) to the same teacher for multiple years in a row. Education researchers have documented benefits of looping. For example, Cistone and Shneyderman (2004) compare average achievement of students in schools with looping to those without and find that students perform better in the looping schools. Our findings corroborate and build on their results; our data allow us to not only strip away more general school fixed effects, but also student and teacher fixed effects. We therefore identify a very clean estimate of the impact of repeat student-teacher matches. And, given the estimated effect is positive, we think that our paper motivates looping as a beneficial and relatively low-cost policy that should be given due consideration.

REFERENCES

Chetty, Raj, et al. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star." *Quarterly Journal of Economics* 126.4 (2011): 1593-1660.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates." *American Economic Review* 104.9 (2014): 2593-2632.

Cistone, Peter, and Aleksandr Shneyderman. "Looping: An Empirical Evaluation." *International Journal of Educational Policy, Research, and Practice: Reconceptualizing Childhood Studies* 5.1 (2004): 47-61.

Dee, Thomas S. "A teacher like me: Does race, ethnicity, or gender matter?" *American Economic Review* 95.2 (2005): 158-165.

Egalite, Anna J., Brian Kisida, and Marcus A. Winters. "Representation in the classroom: The effect of own-race teachers on student achievement." *Economics of Education Review* 45 (2015): 44-52.

Eide, Eric R., and Mark H. Showalter. "The effect of grade retention on educational and labor market outcomes." *Economics of Education Review* 20.6 (2001): 563-576.

Fairlie, Robert W., Florian Hoffmann, and Philip Oreopoulos. "A community college instructor like me: Race and ethnicity interactions in the classroom." *American Economic Review* 104.8 (2014): 2567-2591.

Gershenson, Seth, and Laura Langbein. "The effect of primary school size on academic achievement." *Educational Evaluation and Policy Analysis* 37.1 (2015): 135S-155S.

Gershenson, Seth, Stephen B. Holt, and Nicholas W. Papageorge. "Who believes in me? The effect of student–teacher demographic match on teacher expectations." *Economics of Education Review* 52 (2016): 209-224.

Hill, Andrew J. "The costs of failure: Negative externalities in high school course repetition." *Economics of Education Review* 43 (2014): 91-105.

Hill, Andrew J., and Daniel B. Jones. "Rosenthal Revisited: Self-Fulfilling Prophecies in the Classroom." Working Paper.

Hoffmann, Florian, and Philip Oreopoulos. "A professor like me the influence of instructor gender on college achievement." *Journal of Human Resources* 44.2 (2009): 479-494.

Jackson, C. Kirabo. "Teacher Quality at the High-School Level: The Importance of Accounting for Tracks." *Journal of Labor Economics* (2014).

Jacob, Brian A., and Lars Lefgren. "The effect of grade retention on high school completion." *American Economic Journal: Applied Economics* 1.3 (2009): 33-58.

Kuziemko, Ilyana. "Using shocks to school enrollment to estimate the effect of school size on student achievement." *Economics of Education Review* 25.1 (2006): 63-75.

Manacorda, Marco. "The cost of grade retention." *Review of Economics and Statistics* 94.2 (2012): 596-606.

Papageorge, Nicholas W., Seth Gershenson, and Kyungmin Kang. "Teacher Expectations Matter." *No. 10165. Institute for the Study of Labor (IZA)* Working Paper (2016).

Rothstein, Jesse. "Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement." *Quarterly Journal of Economics* 125.1 (2010): 175-214.

Tables

Table 1: Descriptive statistics: students and teachers in NC, 3rd to 5th grade, math, 1997-2013

	Unique students		Unique teachers		Student-grades	
	Mean	(s.d.)	Mean	(s.d.)	Mean	(s.d.)
Female	0.493	(0.500)	0.914	(0.280)		
Minority	0.403	(0.490)	0.193	(0.290)		
Black	0.273	(0.445)	0.129	(0.335)		
Ever repeat student-teacher match	0.030	(0.169)	0.094	(0.293)		
Teacher z-score			0.022	(0.785)		
Teacher years of experience			9.084	(9.238)		
Repeating student-teacher match					0.012	(0.111)
Number of teachers in grade (same school)					4.695	(2.136)
Class size					24.02	(11.11)
Indicator for "looping class": share of repeating student-teacher matches in class:						
> 0.2					0.022	(0.148)
> 0.4					0.012	(0.109)
> 0.6					0.007	(0.086)
> 0.8					0.003	(0.059)
Observations	2,111,082		70,471		5,135,946	

Table 2. Effect of repeat student-teacher matches on standardized test scores in 3rd to 5th grade

	(1)	(2)	(3)	(4)	(5)	(6)
Repeating student-teacher match	0.123***	0.149***	0.129***	0.036***	0.024***	0.021***
	(0.015)	(0.011)	(0.006)	(0.005)	(0.004)	(0.004)
School fixed effects	N	Y	N	N	N	Y
Teacher fixed effects	N	N	Y	N	Y	Y
Student fixed effects	N	N	N	Y	Y	Y
Observations	5,122,520	5,122,520	5,122,520	5,122,520	4,689,819	4,689,800
R-squared	0.000	0.114	0.181	0.895	0.902	0.903

All regressions include year and grade fixed effects. Robust standard errors clustered at school level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3. Potential selection in repeat student-teacher matches and other robustness checks

	Potential selection				Other robustness checks	
	(1)	(2)	(3)	(4)	(5)	(6)
Ever repeat student-teacher match with current teacher	0.108***	0.133***	0.017***	0.005		
	(0.014)	(0.006)	(0.006)	(0.004)		
Repeating student-teacher match	0.017***	0.012***	0.028***	0.022***	0.024***	0.022***
	(0.006)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)
Number of teachers in grade fixed effects	N	N	N	N	Y	N
Class size fixed effects	N	N	N	N	N	Y
Teacher fixed effects	N	Y	N	Y	Y	Y
Student fixed effects	N	N	Y	Y	Y	Y
Observations	4,689,819	4,689,819	4,689,819	4,689,819	4,689,819	4,689,819
R-squared	0.000	0.181	0.895	0.902	0.902	0.902

All regressions include year and grade fixed effects. Robust standard errors clustered at school level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4. Heterogeneity by teacher quality

	(1)	(2)	(3)	(4)
Repeating student-teacher match	0.011**	0.021**	0.026***	0.033***
	(0.005)	(0.010)	(0.004)	(0.007)
Repeating student-teacher match X Teacher in 2nd VAM quartile		-0.013		
		(0.014)		
Repeating student-teacher match X Teacher in 3rd VAM quartile		-0.012		
		(0.014)		
Repeating student-teacher match X Teacher in 4th VAM quartile		-0.016		
		(0.014)		
Prev. teacher in 2nd VAM quartile		0.010***		
		(0.002)		
Prev. teacher in 3rd VAM quartile		0.019***		
		(0.002)		
Prev. teacher in 4th VAM quartile		0.024***		
		(0.002)		
Repeating student-teacher match X Teacher z-score			0.008	
			(0.005)	
Repeating student-teacher match X Teacher years of experience				-0.001
				(0.000)
Teacher fixed effects	Y	Y	Y	Y
Student fixed effects	Y	Y	Y	Y
Observations	2,190,826	2,190,826	3,350,348	3,352,679
R-squared	0.929	0.929	0.908	0.908

All regressions include year and grade fixed effects. Robust standard errors clustered at school level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5. Heterogeneity by student race

	(1)	(2)	(3)	(4)
Effects on non-minority students:				
Repeating student-teacher match	0.017***			
	(0.004)			
Effects on minority students:				
Repeating student-teacher match	0.034***			
	(0.006)			
Effects on white students:				
Repeating student-teacher match		0.019***		-0.013
		(0.005)		(0.015)
Student-teacher race match				0.006**
				(0.003)
Repeating student-teacher race match X Student-teacher race match				0.034**
				(0.016)
Effects on black students:				
Repeating student-teacher match		0.034***		0.033***
		(0.006)		(0.007)
Student-teacher race match				0.012***
				(0.004)
Repeating student-teacher race match X Student-teacher race match				0.003
				(0.014)
Other effects:				
Repeating student-teacher match			0.028***	
			(0.006)	
Student-teacher race match			0.007***	
			(0.001)	
Repeating student-teacher match X Student-teacher race match			-0.007	
			(0.007)	
Repeating student-teacher match X Female student				
Teacher fixed effects	Y	Y	Y	Y
Student fixed effects	Y	Y	Y	Y
Observations	4,689,797	4,052,682	4,689,797	4,052,682
R-squared	0.902	0.903	0.902	0.903

All regressions include year and grade fixed effects. Robust standard errors clustered at school level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6. Class spillovers in effect of repeat student-teacher matches

	(1)	(2)	(3)	(4)
Share x (see footnote)	*0.2*	*0.4*	*0.6*	*0.8*
Panel A:				
Looping class	0.015***	0.029***	0.030***	0.033***
	(0.004)	(0.005)	(0.006)	(0.009)
Panel B:				
Repeating student-teacher match	0.021***	0.011***	0.016***	0.020***
	(0.004)	(0.004)	(0.004)	(0.004)
Looping class	0.004	0.022***	0.018***	0.016*
	(0.004)	(0.006)	(0.007)	(0.009)
Panel C:				
Repeating student-teacher match	0.008	0.006	0.015***	0.020***
	(0.008)	(0.005)	(0.005)	(0.004)
Looping class	0.002	0.015**	0.016*	0.009
	(0.005)	(0.007)	(0.008)	(0.015)
Repeating student-teacher match X Looping class	0.016*	0.014*	0.003	0.008
	(0.009)	(0.008)	(0.010)	(0.016)
Teacher fixed effects	Y	Y	Y	Y
Student fixed effects	Y	Y	Y	Y
Observations	4,689,819	4,689,819	4,689,819	4,689,819

"Looping class" defined as class in which share of repeating student-teacher matches exceeds x, where x is stated in first row. All regressions include year and grade fixed effects. Robust standard errors clustered at school level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

APPENDIX

Appendix Tables

Appendix Table 1. Effect of repeat student-teacher matches on standardized language test scores

	(1)	(2)	(3)	(4)	(5)	(6)
Repeating student-teacher match	0.085***	0.098***	0.104***	0.024***	0.013***	0.012***
	(0.013)	(0.009)	(0.005)	(0.004)	(0.003)	(0.003)
School fixed effects	N	Y	N	N	N	Y
Teacher fixed effects	N	N	Y	N	Y	Y
Student fixed effects	N	N	N	Y	Y	Y
Observations	5,089,650	5,089,650	5,089,650	5,089,650	4,654,251	4,654,234
R-squared	0.000	0.106	0.153	0.885	0.880	0.881

All regressions include year and grade fixed effects. Robust standard errors clustered at school level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2. Effect of repeat sequential-grade student-teacher matches

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A:						
Repeating student-teacher match in sequential grade	0.124***	0.150***	0.133***	0.039***	0.027***	0.025***
	(0.016)	(0.012)	(0.006)	(0.005)	(0.004)	(0.004)
Panel B:						
Repeating student-teacher match	0.117***	0.139***	0.090***	0.008	-0.003	-0.007
	(0.024)	(0.020)	(0.014)	(0.012)	(0.010)	(0.010)
Repeating student-teacher match X Sequential grade	0.007	0.011	0.043***	0.031**	0.030***	0.032***
	(0.027)	(0.022)	(0.015)	(0.013)	(0.011)	(0.011)
School fixed effects	N	Y	N	N	N	Y
Teacher fixed effects	N	N	Y	N	Y	Y
Student fixed effects	N	N	N	Y	Y	Y
Observations	5,122,520	5,122,520	5,122,520	5,122,520	4,689,819	4,689,800

All regressions include year and grade fixed effects. Robust standard errors clustered at school level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table 3 (alternate version of Table 5). Heterogeneity by student race

	(1)	(2)	(3)	(4)
Repeating student-teacher match	0.017***	0.019***	0.028***	-0.013
	(0.004)	(0.005)	(0.006)	(0.015)
Repeating student-teacher match X Minority student	0.017***			
	(0.006)			
Repeating student-teacher match X Black student		0.015**		0.046***
		(0.007)		(0.016)
Student-teacher race match			0.007***	0.006**
			(0.001)	(0.003)
Repeating student-teacher match X Student-teacher race match			-0.007	0.034**
			(0.007)	(0.016)
Student-teacher race match X Black student				0.007
				(0.006)
Repeating student-teacher match X Student-teacher race match X Black student				-0.031
				(0.022)
Repeating student-teacher match X Female student				
Teacher fixed effects	Y	Y	Y	Y
Student fixed effects	Y	Y	Y	Y
Observations	4,689,797	4,052,682	4,689,797	4,052,682
R-squared	0.902	0.903	0.902	0.903

All regressions include year and grade fixed effects. Robust standard errors clustered at school level in parentheses. *** p<0.01, ** p<0.05, * p<0.1