Please do not cite without authors' permission.

School Quality and the Black-White Achievement Gap

Eric A. Hanushek and Steven G. Rivkin *

December 2005
Revised June 2006

ABSTRACT

Substantial uncertainty exists regarding the impact of school quality on the black/achievement gap. In this paper we use data from both the Texas Schools Project (TSP) panel data and the Early Childhood Longitudinal Survey (ECLS) to investigate the contribution of schools to the racial achievement gap in both elementary and middle school. Following a decomposition of changes in the achievement gap into between- and within-school components, we investigate the effects of specific school variables including teacher experience, student turnover and student racial composition on the black/white differential. One issue that we examine in some detail is the possibility that changes in the racial achievement gap and the effects of specific school characteristics vary by initial achievement and by gender. Our results differ noticeably from the other recent analyses of the black-white achievement gap by providing strong evidence that school quality has a substantial effect on the differential. We find that the majority of the expansion of the achievement gap with age occurs within rather than between schools and the aforementioned school characteristics explains much of the growth in the achievement gap.

*Stanford University, University of Texas at Dallas, and National Bureau of Economic Research; and Amherst College, University of Texas at Dallas, and National Bureau of Economic Research, respectively.
Support for this work has been provided by the Packard Humanities Institute.

Schools, Peers, and the Black-White Achievement Gap

By Eric A. Hanushek and Steven G. Rivkin

1. Introduction

Differences in cognitive skills have been shown to explain the bulk of the gap between blacks and whites in lower school attainment and in wages, motivating aggressive policies to raise the quality of education for blacks.[1] The landmark decision in *Brown v Board of Education* that attacked racial segregation of schools was the modern beginning of concerted federal, state, and local actions directed at improving black achievement.[2] Along with subsequent court cases, *Brown* ushered in a profound change in both school and peer characteristics, while contemporaneous increases in school spending, brought on in part by school finance litigation, further raised the amount of resources devoted to black students in the public schools. Yet the stubborn persistence of racial achievement gaps coupled with various strands of recent research has led to questions about the power of schools to improve academic, social, and economic outcomes.

Evidence on recent trends in the racial test score gap raises additional doubts about the efficacy of these efforts to reduce the achievement differential. Earlier optimism about narrowing gaps (Jencks and Phillips (1998)) has dissipated with new evidence that the black-white achievement gap stayed constant or even grew during the 1990s (National Center for Education Statistics (2005)). Although the dramatic desegregation, smaller classes, and greater school resources that have been introduced by policy have strong popular appeal, substantial uncertainty remains over the actual effects on the quality of

[1] O'Neill (1990) and Neal and Johnson (1996) provide evidence on wage differences, and Rivkin (1995) provides evidence on differences in educational attainment.
[2] *Brown v. Board of Education*, 347 U.S. 483 (1954).

instruction and student outcomes over both the short and long term. Accumulated evidence does not provide strong support for the belief that higher expenditures typically leads to substantial improvements in the quality of instruction.[3] Although research provides more consistent evidence supporting a positive relationship between black student achievement and the proportion of white students in a school, direct evidence on the benefits of school desegregation is also rather limited.[4]

This conundrum is highlighted by recent work of Fryer and Levitt (2004, (2005). They find, unsurprisingly, that a substantial racial achievement gap exists at entry to school and increases with age. However, unlike the preschool gap, the increasing gap is not readily explained by family background factors, and the majority of the increase occurs within rather than between schools. Moreover, specific school factors appear to account for little of the increase. Nonetheless, reinforcing the confusion in the relevant research, Murnane, Willett, Bub, and McCartney (2005) cannot replicate either the basic school patterns of the achievement gap or the influence of measured family background on the gaps when they go to a different, but in some ways richer, data base.

This paper investigates the contribution of schools to the racial achievement gap as it evolves from kindergarten to the end of middle school. Following a decomposition of changes in the achievement gap into between- and within- school components, we investigate the effects of specific school variables including teacher experience, student turnover and student racial composition on the black/white differential. These factors not only have been shown to affect achievement but also are highly uneven in their incidence

[3] See Hanushek (2003) for a summary of the evidence.
[4] Review of the evidence surrounding desegregation actions provides limited support for positive achievement effects Schofield (1995). Guryan (2004) finds that desegregation reduced the probability of

by race.[5]

We also examine in some detail the possibility that changes in the racial achievement gap vary by initial achievement. This investigation has both analytical and policy importance. Differences in the early achievement distributions for blacks and whites complicate comparisons by race, particularly if growth rates differ systematically by initial achievement either due to actual differences in skill acquisition or limitations in the measurement of achievement. In addition, several hypotheses have been offered that suggest that the gap may grow more rapidly for initially high achieving blacks. On the one hand, blacks who excel in the early grades may face the strongest peer pressure against academic success. Alternatively, higher achieving blacks may fall further from the center of their school's achievement distribution and be less likely to participate in an academic program that facilitates continued excellence.[6] Importantly, we consider the effects of test measurement error and related regression to the mean on the pattern of racial achievement differences.

The analysis traces the black-white achievement gap from kindergarten through grade 8. We use data from the Early Childhood Longitudinal Survey (ECLS) – the basis for the Fryer and Levitt work – for analysis through grade 5 and the Texas Schools Project (TSP) panel data for grades 3 through 8. While the TSP data are richer and more extensive and offer the clearest picture of school influences, they do not provide student observations

---

dropping out of high school, though data limitations and methodological concerns raise questions about the findings.

[5] Hanushek, Kain, and Rivkin (2004a) investigate the effects of student mobility, Rivkin, Hanushek, and Kain (2005) investigate the effects of teacher experience, and Hanushek, Kain, and Rivkin (2006) investigate the effects of racial composition.

[6] Fryer and Levitt (2005) consider a related hypothesis through comparing performances of blacks and whites on alternative cognitive tests and suggest that blacks may indeed be doing more poorly on tests of higher order skills. Murnane, Willett, Bub, and McCartney (2005) further question the impact of test content and score calculations on the pattern of achievement gaps.

in the earliest grades.

Our results differ sharply from the other recent analyses of the black-white achievement gap. First, we find that the majority of the expansion of the achievement gap with age occurs between rather than within schools in both the ECLS and TSP data. The contrast with the findings of Fryer and Levitt (2005) appears to result from a problem with their achievement decomposition. Second, we find that identifiable school factors – the rate of student turnover, the proportion of teachers with little or no experience, and student racial composition explains much of the growth in the achievement gap between grades 3 and 8. Unfortunately, the structure of the ECLS does not permit the estimation of the causal effects of these variables for the grades and test instruments in that sample. Nonetheless, the similar race differences in these school and peer characteristics and the much larger increases in the between-school component of the racial achievement gap in the early ECLS grades suggests the impact of schools is likely to be as large if not larger in the earlier grades.

The next section describes the ECLS and TSP data sets used in this analysis. Section 3 describes changes in the racial achievement gap with age for all blacks and whites and by initial achievement and gender. This section also decomposes the gap into within-and between-school components to illustrate the potential importance of schools in explaining growth in the achievement differential as students progress through school. Section 4 describes the empirical model prior and then examines the effects of specific school factors on achievement along with their contribution to racial achievement differences. The final section summarizes the findings and discusses potential implications for policy.

2. ECLS and TSP Data

This paper employs both the Early Childhood Longitudinal Survey Kindergarten Cohort (ECLS) and Texas Schools Project (TSP) data sets in the investigation of the black/white achievement gap. The ECLS is a nationally representative sample of achievement data for grades K-5, and the TSP data contain administrative information on the universe of Texas public school students for grades 3-8. Together these data span the elementary school years, and the stacked panels contained in the TSP data facilitate the estimation of school and peer group effects on achievement.

A. ECLS Data

The ECLS is a survey of the National Center for Education Statistics that is designed to provide extensive information on the early elementary school years. To date six waves of data have been collected, beginning with the base year kindergarten survey in fall 1998. Follow up surveys were completed in the spring of kindergarten, the spring of the subsequent academic year and every two years thereafter for all students and in the fall of the year following kindergarten for a much smaller sub-sample. Students remaining with their cohort were thus surveyed twice in kindergarten, once or twice in first grade, in third grade, and in fifth grade. Because the fall, first grade survey was administered to only a subset of students, we do not use information from that wave.

Standardized mathematics and reading tests were administered in each of the waves along with child surveys that elicit information on race, ethnicity, family financial circumstances, parental education and employment, and a number of other variables. Information on teacher, school, and student demographics was also collected from schools

and teachers each academic year, and sampling weights were provided in order to make the data nationally representative. Because of the relatively small sample size, limited and potentially noisy information on family background, and concerns about the possibility of omitted variables bias that cannot be mitigated using the panel data techniques used in the analysis of the TSP data, we do not use the ECLS data in the estimation of school, peer, and teacher effects. Rather we use the test scores to describe the evolution of the racial achievement gap for this cohort and use the information on teachers and peers to characterize racial differences in the school environment.

A two stage adaptive testing procedure was used to measure achievement. Students first completed a short pretest that sorted them into categories on the basis of the number correct. Students were administered different tests depending upon the pretest score, and test administrators used item response theory algorithms to grade the examinations. Theoretically the tests are vertically scaled such that a given point differential reflects a given difference in knowledge throughout the scale regardless of whether the differential reflects different scores on the same test or results for different grades.

A potentially serious problem is the high share of students lacking test information in a given year. Among those who participate in the five waves used in this paper, roughly 50 percent of blacks and over one third of whites do not have test scores.

B. TSP Data

The TSP data set is a unique stacked panel of school administrative data constructed by the UTD Texas Schools Project. The data we employ track the universe of Texas public elementary students as they progress through school. For each cohort there are over 200,000 students in over 3,000 public schools. Unlike many data sets that sample

only small numbers from each school, these data enable us to create quite accurate measures of peer group characteristics. We use data on four cohorts for grades three (earliest grade tested) through eight. The most recent cohort attended 8th grade in 2002, while the earliest cohort attended 8th grade in 1999.

The student data contain a limited number of student, family, and program characteristics including race, ethnicity, gender, and eligibility for a free or reduced price lunch (the measure of economic disadvantage). The panel feature of the data, however, is exploited to account implicitly for a more extensive set of background characteristics through the use of a value added framework that controls for prior achievement. Importantly, students who switch schools can be followed as long as they remain in a Texas public school.

Beginning in 1993, the Texas Assessment of Academic Skills (TAAS) was administered each spring to eligible students enrolled in grades three through eight. The tests, labeled criteria referenced tests, evaluate student mastery of grade-specific subject matter. This paper presents results for mathematics. Each math test contains approximately 50 questions. Because the number of questions and average percent right varies across time and grades, we transform all test results into standardized scores with a mean of zero and variance equal to one, which transforms the outcome into a measure of relative position in the achievement distribution.

Notice that the persistence of a constant differential in terms of relative score does not imply a constant knowledge gap. If the variance in knowledge grows with age and time in school, as we believe most likely, any deterioration in the relative standing of blacks on the achievement tests would understate the increase in knowledge inequality.

Importantly, the student database can be linked to information on teachers and schools. The school data contain detailed information on individual teachers including grade and subject taught, class size, years of experience, highest degree earned, race, gender, and student population served. While individual student-teacher matches are not possible, students and teachers can be uniquely related to a grade on each campus. Each student is assigned the average class size and the distribution of teacher characteristics for teachers in regular classrooms for the appropriate grade, school, subject, and year.

3. Racial Achievement Gaps

Beginning with the "Coleman report," *Equality of Educational Opportunity* (Coleman et al. (1966)), test score decompositions have been used to learn about the contribution of schools to the variation in achievement. The logic is simply that school policies mainly affect differences across schools, so a small proportion of achievement variance lying between schools is suggestive of a limited role of schools as opposed to families and other factors that vary with the individual child. There are of course reasons that we return to below about why this rough logic might fail, but it is useful to understand the basic pattern of achievement gaps.[7]

Fryer and Levitt (2004, 2005) report that most the average black/white achievement differences lie within schools, but their approach fails to capture accurately the contributions of the within- and between-school components. Although there is a simple decomposition of the variance of achievement into between- and within-school

[7] Sorting by families and teachers leads to systematic differences in family and community background among schools that is correlated to school and teacher quality, and differences in the quality of instruction exist within schools (see, for example, Rivkin, Hanushek, and Kain (2005)).

components, this calculation does not carry over to consideration of the mean achievement gap. Specifically, the between school component of the mean gap does not equal the average overall gap minus the student weighted average within school gap (the coefficient on black from a school fixed effect achievement regression) except in the special and uninteresting case in which the enrollment shares of blacks and whites are identical in all schools. Given the uneven distribution of whites and blacks among schools, this calculation produces erroneous measures of the within and between school contributions by ignoring the effects of the distribution of blacks and whites among schools.

An example illustrates the problem with setting the between school component equal to the overall gap minus the fixed effect coefficient. Consider a sample of 1,000 schools where all but one is completely segregated. The fixed effect estimator of the "within" component of the achievement gap would come entirely from the single school containing both black and white students. This one school has virtually no effect on the overall gap and no information on the difference between the segregated black and white schools. Nonetheless, if the within school difference in this one school happened to be larger than the overall gap, this approach would imply that the between school contribution was negative and that differences among schools actually reduced the overall achievement gap.

Equation (1) shows the appropriate weighting of the two components such that the between-school component (first term in brackets) and the within-school component (second term in brackets) adds up to the average difference in white and average black achievement scores ($\overline{A}_w - \overline{A}_b$).[8] The first term shows explicitly how differences in the

[8] The appendix provides a derivation of this decomposition.

distributions of blacks and whites among schools with different average levels of achievement determines the between-school component, where n_{ds}/n_d is the share of demographic group d in school s.[9]

$$\overline{A}_w - \overline{A}_b = [\sum_s \frac{n_{ws}}{n_w}\overline{A}_s - \sum_s \frac{n_{bs}}{n_b}\overline{A}_s] + [\left(\frac{1}{n_w} + \frac{1}{n_b}\right)\sum_s (\overline{A}_{ws} - \overline{A}_{bs})\alpha_s(1-\alpha_s)n_s]$$

The second term is a weighted average of black/white achievement differences within each school, where α_s equals the proportion of black and white students in school s who are black and 1- α_s equals the proportion who are white. Notice that the weight for school s is neither proportional to the school's share of total enrollment of blacks or whites or its share of enrollment in schools with both blacks and whites (the case with a school fixed effects estimate). Rather the weight for school s depends non-linearly on both total enrollment and the distribution by race: schools with higher total enrollment and more equal enrollment shares receive greater weight. Notice that equation (1) makes clear that the within school contribution in the above example with the single integrated school approaches zero, meaning that essentially all of the differential is attributed to the between school component.

A. Average Black/White Achievement Differences

Tables 1 and 2 report the evolution of the black/white achievement gap in the ECLS and TSP data, respectively, and decompose that gap into within- and between-school components. Table 1 reports the pattern of achievement gaps for the ECLS using the sample of students with a complete set of test scores and the sampling weights relevant for

[9] n_b is the total number of black students, and n_{bs} is the number of black students in school s, with parallel definitions for white students (w).

students who participated in the five waves used in this paper.[10] The overall gap begins at 5.4 points and increases by 2.5 points during kindergarten, 4.4 points during first grade, 5.7 points during grades two and three, and by 1.4 during grades four and five. Although the gap continues to widen through fifth grade, the pace slows following first grade. The much high rate of missing tests for blacks than whites suggests that these trends provide a lower bound on the average achievement gap for all blacks and whites and likely also on the growth in that gap with age.

Virtually all of the grade to grade increases in the overall gap occur between schools. The increase in the between-school component accounts for 12.8 points out of the total increase of 14 points throughout the period. But, if instead we estimated the within-school component by the black coefficient from the school fixed effect regression, it would suggest an increase of 11 points, accounting for almost 80 percent of the overall gap during this period.[11]

Table 2 reports mean scores for grades 3, 5, and 8 from the TSP data calculated over all test-takers and over those who remain with their initial cohort for all grades.[12] (Note that these comparisons are calculated in terms of standard deviations, because no IRT scores are available for the Texas tests).[13] A comparison of the gaps produced by the repeated cross sections (right panel) and sample of students with a complete set of test observations who progress with their class (left panel) highlights the importance of grade

---

[10] Scores are scaled according to item response theory (IRT), permitting them to be equated across grades. These calculations use the sampling weights from survey (where the decomposition follows equation 1 based on weighted student counts for each school). Calculations not using the sampling weights are very similar in both magnitude and pattern.

[11] This is very similar to the pattern estimated by Fryer and Levitt.

[12] We do not report all grades, because movement from elementary to junior high or middle school produces a great deal of temporary test volatility in grades 6 and 7 that disappears within a year following the transition..

retention, special program assignment, and other factors that determine test taking patterns. As Appendix Table A1 shows, grade retention and special education classification rates for blacks exceed those for whites in all grades, and the gaps are higher in grades 5 and 8. Because test taking rates are far lower for children classified as special needs, a larger share of black 3rd grade test takers will either lack a test observation or no longer attend school with their original cohort by the end of 8th grade.

The probability of grade retention and of special education classification is higher for those in the lower portion of the achievement distribution, implying that blacks in the complete cohort sample will tend be a more select group than whites in the complete cohort sample. Across grades, the sample of blacks will tend to have been in school longer on average than whites because of grade retention, and this works to narrow the black-white gap when taken as a snapshot at a given grade level. Both the lower 3rd grade test score gap and larger increase in the gap between 3rd and 8th grade for the complete cohort sample result from these differences in grade retention and special education classification.

In the complete panel sample, the gap rises from 0.59 standard deviations at the end of 3rd grade to 0.65 in 5th grade and to 0.70 in 8th grade. The pattern of achievement gap change is similar to that observed in the ECLS. Not only does the yearly rate of increase diminish with age, but the between-school component accounts for roughly 75 percent of the change (0.9 versus 0.03).

Because families sort purposefully into communities and school districts, family, neighborhood, and other non-school factors likely contribute to the between-school increase in the black/white achievement gap. Nevertheless, the patterns observed in Tables

---

[13] Murnane, Willett, Bub, and McCartney (2005) suggest that the different units of analysis can affect the results, but we have no way to deal with that issue here.

1 and 2 leave open the possibility that school quality accounts for much of the rise in the elementary school racial achievement gap.

B. Racial Differences by Initial Achievement

By dividing the 3rd grade achievement distribution into a number of equal sized intervals, systematic variation in the change in the achievement gap between 3rd and 8th grades by initial achievement can be identified. This is complicated, however, because test scores measure actual knowledge with error, leading to misplacement of students into categories and biasing the calculation of gaps across grades. If errors are uncorrelated over time, those initially placed in high achievement categories will tend to draw less positive errors in the subsequent year, while those placed in the lower categories will tend to draw more positive errors in the subsequent year. Therefore regression to the mean will account for a portion of the observed difference in test score changes across categories.

Such regression to the mean also complicates black-white comparisons by achievement level because of differences in the actual initial skill distribution. Table 3 illustrates the general problem. Consider the true skill distributions (which has three achievement categories) shown in the top panel of Table 3 for blacks and whites, where the distribution for blacks is more concentrated in the lower categories than the distribution for whites.

The bottom panel introduces misclassification that comes through measurement errors in the tests. P_{ij} is the probability that somebody whose true ability is category i but the student is observed in category j (e.g., P_{LH} is the probability that a low ability person will have a test score in the high ability category). Even when the measurement error is randomly distributed with no differences by race, the pattern of achievement by race will

be distorted. More whites than blacks have been misclassified into the lowest observed skill category, while the opposite is true for the highest observed skill category. Strictly as a result of this regression to the mean phenomenon, the expected achievement gain in the next period is higher for whites than for blacks throughout the initial skill distribution if the errors leading to misclassification are uncorrelated over time.

The pattern of observations illustrated in Table 3 invalidates the simple categorization of students on the basis of initial test scores. To overcome this problem, we use a test in a different subject to categorize students by initial mathematics skill level, based on the assumptions of positive correlations across subjects in true skill and of no correlation in the test measurement errors across subjects. This scheme severs the link between initial category and expected difference in the error realizations for the initial and subsequent periods.

Tables 4 reports the overall and between school components of the black/white achievement gap in mathematics for the spring of kindergarten, first, third and fifth grade using the ECLS sample by six initial reading achievement categories.[14] Other than the highest category where the gap declines following 1st grade and in the period as a whole, Table 4 illustrates that the gap in the ECLS sample increases much more prior to 3rd grade than between 3rd and 5th grade. Although the largest increase occurs in the bottom reading category, the overall increase in these five categories does not vary much by initial reading achievement. Finally, the bulk of the overall changes in all categories results from between-rather than within-school changes.

Similar to the national trends from the ECLS for grades k thru 5, Figure 1 (derived

[14] We use the spring kindergarten rather than the fall kindergarten test to divide students in the ECLS because of the lack of dispersion in the fall kindergarten reading test distribution.

from Table 5) shows that growth in the achievement gap for the TSP sample tended to be noticeably larger in the earlier grades (3rd to 5th). However, in contrast to the pattern observed in Table 4, Figure 1 reveals a pronounced ordering in the magnitude of change by initial achievement categories, particularly following 5th grade (solid bars). Between grades 5 and 8 the overall gap increased in only one of the seven bottom categories and by only 0.01 standard deviations, while the gap increased by at least 0.05 standard deviations in all but one of the next six groups and by at least 0.09 standard deviations in the top three groups. For the period as whole, the gap increased by at least 0.25 standard deviations in the top three groups, between 0.17 and 0.21 in five of the next 6 groups, and by less than 0.10 in the seven lowest categories.

Similar to the pattern observed in the ECLS sample, Figure 2 shows that between-school changes account for the bulk of the overall changes observed between 3rd and 5th grade in Texas across the initial reading achievement distribution. However, following 5th grade there is a marked divergence in the impacts of the between-school component by 3rd grade reading category. The between-school gap actually declined in each of the bottom seven categories, and in six of the seven this decline led to a reduction in the overall differential despite offsetting changes in the within-school component in most. In contrast, the between-school gap did not fall in any of the remaining groups and tended to move in the same direction as the within-school changes. Consequently the overall gap increased for students in the upper reading achievement categories, though as noted the magnitude of the gap expansion following 5th grade tended to be much smaller than the changes observed between grades 3 and 5.

C. Differences by Gender

The recent widening of the gender gap in high school and college completion, particularly pronounced for blacks, raises the possibility of a gender differential in the evolution of the elementary school achievement gap. Appendix Tables a2 and a3 report overall gender differences for the ECLS and TSP samples, respectively, and Tables a4 and a5 report gender differences by initial reading achievement.

All in all, the tables provide little or no evidence of differential patterns by gender, and particularly no evidence of a sharp drop in the achievement of black boys during these grades. This holds across the initial reading achievement distribution and for both total changes and the within-and between-school components.

It is important to recognize, however, that these figures come from samples with complete sets of test score observations and in the case of the TSP results with samples restricted to those who remain with their cohort. Differential rates of special education classification, grade retention, and other sources of missing tests could conceal substantial differences in average achievement growth for demographic groups as a whole. More importantly, they could hide serious education problems for students at the lower end of the achievement distribution.

Appendix Table a1 reports the distribution of test taking status by gender and race in the TSP sample and shows that black boys are roughly twice as likely as white boys and almost twice as likely as black girls to have missing test scores in a given year, primarily due to a much higher rate of special education classification. Consequently the achievement comparisons based on the TSP sample do not capture the educational difficulties of many black boys, particularly those at the lower end of the initial achievement distribution. By comparison, there is little gender differential in the ECLS

sample rates of test taking for blacks and whites.

4. School and Peer Effects on Achievement Gaps

This section investigates the extent to which specific teacher and school variables account for the growth in the achievement gap during the school years and differences by initial achievement. Although some recent studies including Fryer and Levitt (2004) have not found that school factors account for much if any of the growth in the achievement gap, these results are inconsistent with other research that highlights the significant effects of specific school and peer factors that clearly differ by race.

Our goal here is to assess whether schools have a discernible impact on the growth of the racial achievement gap. We pursue a conservative strategy that concentrates on that portion of the achievement variance that can be credibly related to the causal influence of specific school factors. As such we eliminate systematic components of schools that operate on the achievement gap where we cannot be certain about their causal impact. In fact, because the ECLS data offer virtually no chance of credibly identifying the causal impact of any of the school factors, we limit our analysis just to the Texas data. This is unfortunate because the Texas analysis begins with grade 4 and thus does not include the earlier grades where there are larger changes in achievement gaps.

In order for a specific variable to contribute to the achievement gap, it must be a significant determinant of achievement and the distribution of the variable must differ systematically by race. Prior research leads us to focus on a small number of variables shown to be strongly related to achievement. Our analytical plan is first to obtain estimates of their causal effects and then to use these estimates in combination with information

about the racial distributions of these variables to estimate their contributions to the racial achievement gaps described in the previous section.

A. Empirical Model

The central problem in modeling the impact of teacher and school variables is the non-randomness and interdependence of the allocation of students and teachers among schools. Both school resources and racial composition, for example, are the outcomes of decisions made by families, school officials, legislators, and in some cases judges, and these are likely to interrelate with a range of factors that directly and indirectly influence achievement.

The candidates for consideration in this modeling are school factors that have previously been singled out as having a significant and causal impact on achievement. Equation (2) highlights the key identification issues that must be addressed in the absence of random assignment. Here achievement (A) for student i in cohort c in grade G and school s is modeled as a function of student, family, teacher, and peer factors:

(2) $A_{iGs}^{c} = \alpha_{iG}^{c} + \beta X_{iGs}^{c} + \delta T_{iGs}^{c} + \lambda P_{iGs}^{c} + S_{iGs}^{c} + e_{iGs}^{c}$

where P is a vector of peer variables that includes proportion black and proportion of students that are new to the school in grade G, T is a vector of teacher experience variables that includes proportions of teachers in grade G with zero prior years of experience, one prior year of experience, and two prior years of experience, S is class size, X is a vector of student and family background variables that includes indicators for male, black, eligibility for a subsidized lunch, a structural school switch prior to grade g (e.g., a transition from elementary to middle school), and a family initiated school change prior to or during grade g, α is an individual intercept specific to grade G, and e is a stochastic term capturing other

unmeasured influences.[15]

If all variables in T and P were uncorrelated with α and *e,* OLS would yield unbiased estimates of δ and λ. But as noted above, the complications inherent in the determination of peer and teacher characteristics bolstered by existing evidence strongly suggest that typically available controls do not account adequately for potentially confounding factors.

Our basic approach is to include full sets of school-by-grade and school-by-year fixed effects in order to isolate variation in racial composition, teacher experience, and student turnover that is orthogonal to such confounding factors. Not only do these school fixed effects account for all between-school differences including aspects of schools that affect achievement, the school-by-grade terms also eliminate systematic changes across grades in a school common to all years, and the school-by-year terms control for year-to-year changes common to all grades in the school. These school-by-year fixed effects remove in a very general way all variation over time in neighborhood and local economic conditions that likely affect mobility patterns including such things as the introduction of new race-related school policies or the myriad changes documented to occur in "transitional neighborhoods." An economic shock that reduces neighborhood employment and income would not bias the estimates; nor would a shock to local school finances or the quality of the local school board, because each of these would affect all grades in a school. The school-by-grade fixed effects also account for the possibility that achievement trends vary systematically with changes in teacher experience, peer turnover,

[15] All specifications also include a full set of grade-by-year dummy variables to capture changes to the tests. At this point we also eliminate a series of common factors that past work has shown to be unrelated to student achievement, such the amount of teacher graduate education.

or school racial composition as students age.[16]

These fixed effects do not control for any school decisions on classroom placement that might be related to race or student background. For this analysis, we measure teacher and peer variables at the grade rather than classroom level, avoiding the complication introduced by the selective placement of students into classrooms.[17]

In this framework, the remaining variation comes both from mobility induced changes and from persistent cohort-to-cohort differences within schools. Neither mobility nor cohort differences are random, and evidence indicates the mobility affects both movers and other students in the destination school. Moreover, the evidence shows that movers tend to have lower prior achievement, indicating that determinants of learning in prior periods were less conducive to achievement. Therefore, we include indicators for both family initiated and structural moves and control for student differences (captured by α) that may be related to racial composition, turnover, or teacher experience.

Consider Equation (3), which specifies α as a function of prior school and family variables, racial composition in previous grades, and unobserved "ability," γ.[18]

---

[16] Because multiple elementary schools are often combined into a single middle school, blacks in very high percent black elementary schools often attend more racially mixed middle schools. Evidence suggests that these students experience particular difficulties with the transition to middle school which could lead to a spurious positive correlation between achievement and school percent black in the absence of the school-by-grade fixed effects.

[17] While alternative approaches for dealing with classroom placement would be possible, our data do not support classroom specific analysis. Clotfelter, Ladd, and Vigdor (2003) find significant variations in the racial composition of classrooms by district, school, classroom, and academic track in middle school but much less so in primary school. They do not address implications for student performance, but given that the school-by-year and school-by-grade fixed effects account for any persistent placements for a grade and year-to-year school wide changes, such within-school differences should have minimal effect on the estimates in this paper.

[18] Boardman and Murnane (1979) and Todd and Wolpin (2003) also highlight the importance of unobserved ability and the cumulative nature of learning.

$$(3) \qquad \alpha_{iG}^{c} = \beta \sum_{g=1}^{G-1} \theta^{G-g} X_{igs}^{c} + \delta \sum_{g=1}^{G-1} \theta^{G-g} S_{igs}^{c} + \lambda \sum_{g=1}^{G-1} \theta^{G-g} b_{igs}^{c} + (\gamma_{i}^{c} + \lambda \sum_{g=1}^{G-1} \theta^{G-g} \gamma_{i}^{c})$$

This formulation captures the possibility that family, teacher, and peer interactions in grades prior to G establish the knowledge base for learning in grade G and therefore affect achievement at the end of grade G. In a very general manner, the effects of prior period variables are assumed to decline exponentially as a function of time from the present at a constant rate (1-θ), where $0 \le \theta \le 1$.[19]

The term γ captures student differences that remain constant during the schooling years including early childhood influences, prenatal care, heredity, and other systematic factors. Notice that our formulation is learning-based in that the value of γ affects the quantity of skills and knowledge acquired at each grade, and these increments to achievement are subject to depreciation. This explicitly permits the affects of ability on achievement to increase with age.[20]

B. Baseline Results

Table 6 reports estimates of the effects of student turnover, racial composition, and teacher experience on mathematics achievement for three specifications that progressively

[19] At the extreme of θ =0, past inputs are not relevant for current achievement, i.e., having a good fourth grade teacher does not have any implications for math achievement at the end of the fifth grade. On the other hand, θ =1 implies no depreciation of the influence of past inputs, i.e., that the impact of a good fourth grade teacher on 4th grade achievement equals her impact on 5th grade achievement and achievement in all future grades. For convenience, we assume that the effects of prior variables decay at the same rate, although this is not essential for the development below.

[20] The exact formulation and interpretation depends, however, on the measurement of achievement. If measured with vertically integrated tests, differences in γ would contribute to a widening of the skill distribution over time as long as θ were not equal to zero. On the other hand, if skills were measured in distributional terms (as we do here with standardized scores), the complicated final term in parentheses could be replaced with γ_i^c, because ability induced differences in relative achievement would remain constant over time.

add school-by-year and school-by-grade fixed effects. Only black and white non-Hispanic students who remain with their cohort and have non-missing test scores for grades three through eight are included in the sample, and a small number of observations are excluded because of missing information on teachers. There are a total of 1,448,458 observations. All specifications include class size, indicators for economic disadvantage, special education participation, family initiated and structural school changes (all fully interacted by race), and a full set of grade-by-year dummy variables. Preliminary work showed that student proportions Hispanic and Asians were not significantly related to achievement and that their exclusion had virtually no effect on the remaining variable estimates. Absolute values of t-statistics based on robust standard errors clustered by school are reported.

In each instance we investigate whether the specific factor has a direct impact on achievement and then whether there is any indication of a differential impact across black and white students. The results provide strong evidence that a higher proportion of teachers with little or no experience and higher rate of student turnover negatively affect achievement for both blacks and whites, and the estimates suggest that blacks incur a slightly larger cost from having a teacher with no prior experience. On the other hand, a higher proportion of black students decreases achievement for blacks but not for whites.[21] These variable effects are fairly stable across specifications and all significant at the one percent level in the preferred full fixed effects model. Importantly, these factors are correlated with each, suggesting potentially substantial impacts on achievement gaps. Schools with higher turnover and minority enrollment tend to have a higher proportion of inexperienced teachers (Greenberg and McCall (1974), Murnane (1981), Lankford, Loeb,

and Wyckoff (2002), Hanushek, Kain, and Rivkin (2004b)).

In order to assess the impact of these factors on achievement gaps, however, we must combine the estimated models with the distribution of these factors in the population. Consistent with many policy discussions, Table 7 displays the substantial racial differences in the share of teachers with little or no experience, in student turnover, and in student racial composition that are found in both the ECLS and TSP data.[22] Black students are 4-5 percent more likely to have a new teacher and, as is well known, far more likely than whites to attend school with a high black enrollment share. This increased racial concentration averages to a 28 percentage point differential in Texas. Note that the racial composition information in the ECLS data set is much less precise, aggregating all students in schools with 25 percent or more black students into a single category even though 75 percent of black students in the sample fall in this category. Even though the measures are cruder, there is nonetheless clear evidence that black students systematically attend schools with significantly higher black enrollment shares.

The significance of these differential inputs fits together with the significant impacts of each on achievement displayed in Table 6. Together, the findings indicate that these three characteristics may account for a substantial portion of the growth in the racial achievement gap. Table 8 illustrates their respective effects on the annual and cumulative achievement gaps with a simulation that multiplies the changes in characteristics for whites and blacks necessary to eliminate race differences inputs by the respective coefficients. The variable changes for blacks and whites are proportional to the black and white

---

[21] The differential effect of racial composition is consistent with the findings of Hanushek, Kain, and Rivkin (2006) and Angrist and Lang (2004).

[22] Although class size is also frequently mentioned for consideration in the early grades, there are virtually no racial differences in average class size in either the ECLS or TSP data.

enrollment shares, and the simulations use the interaction term coefficients (differential racial impact terms) for all variables. All variable effects are assumed to be constant across grades and to depreciate exponentially at an annual rate of 30 percent (one minus the coefficient on lagged test score).

Importantly, an estimated constant impact across grades and an observed constant racial difference in the factors (say, the prevalence of new teachers) will yield a growing racial gap. The models identify the impact on annual achievement gains, but these accumulate because learning today builds upon the initial knowledge at the beginning of each grade.

Not surprisingly given the size of the black/white differential, the simulation in Table 8 shows that school proportion black accounts for a much larger share of the growth in the gap than student turnover and teacher experience. Together those account for roughly 0.015 standard deviations or almost 20 percent of the increase, while proportion black accounts for 0.033 standard deviations or roughly 40 percent of the increase in the gap between 3rd and 8th grade. The fact that these three characteristics alone appear to account for over half of the growth in the achievement gap provides strong support for the belief that deficiencies related to the school experience inhibit the academic progress of blacks vis-à-vis whites during these grades.

The similarly large black/white differences in teacher experience, student turnover, and proportion black observed in the ECLS data in combination with the much more rapid growth in the achievement gap in the early grades strongly suggests that the impact of

school quality differences is at least as large if not larger in the early elementary grades.[23]

The simulation does make several assumptions that if violated might lead to an overstatement of the actual contribution of these variables. First is the assumption that variable effects remain constant across grades. Although evidence indicates that class size effects diminish with age, there are not similar findings for teacher experience and the student characteristics.[24] However, the slower growth in the between-school component of the racial achievement gap in higher grades suggests that school quality differences may translate into smaller achievement differences at higher grades. If this is in fact the case for these variables, it will reduce the cumulative effect on the achievement gap.

The assumption of a linear effect of proportion black on achievement constitutes a second important assumption, particularly given the magnitude of the black/white difference in average school proportion black. Again, research using the TSP data found little or no evidence of non-linear effects up to a quartic in school proportion black (Hanushek et al, 2006).

B. Differences by initial achievement

An important question is whether these variables can also explain the larger differences observed for students at the higher end of the initial skill distribution. Either variable effects or racial differences in the characteristics could be larger at the upper end of the initial skill distribution. We first consider the pattern of racial differences in the characteristics and then explore the possibility that variable effects may differ systematically by initial achievement.

---

[23] Note, however, that the direct estimates of these factors on early achievement by Murnane, Willett, Bub, and McCartney (2005) does not show their importance for achievement. They further suggest that racial composition is simply a proxy for SES.

[24] Rivkin, Hanushek, and Kain (2005) shows that class size effects fall with age.

Appendix Table A6 provides little or no evidence of systematic differences by initial skill grouping in the racial difference in teacher experience, student turnover, student racial composition, or other school or peer characteristics. In addition, neither the difference in peer average 3rd grade mathematics achievement nor the difference in the share of peers in one of the top four reading groups (not shown) rises along with initial reading score. Thus it does not appear that the pattern of differences in school or peer characteristics accounts for the higher growth in the achievement gap at the higher end of the initial reading test distribution.

However, the pattern of estimates reported in Appendix Table a7 suggests the existence of systematic differences in the effects of racial composition and teacher experience that might account for a portion of the higher increase in the top reading categories. Although the coefficients fluctuate from category to category, the cost of having an inexperienced teacher tends to grow for both whites and blacks as reading achievement declines, while the cost to blacks of a higher proportion black tends to fall. Such a pattern does not emerge for the student turnover variable.

In order to test more formally for systematic differences by reading achievement we divide students into two groups, those in one of the top four categories (high scorers) and the rest, and test for systematic differences in variable effects between the groups. Table 9 reports regression results and F-test p-values for models that fully interact the variables with the indicators for black and high scorers. Supporting the pattern observed in Appendix Table a7, these estimates show that the effects of proportion of teachers with no prior experience are systematically smaller for blacks and whites with higher 3rd grade reading scores, while the adverse effect of a higher proportion black is larger for blacks in

higher reading categories. In terms of magnitudes, the difference in the point estimates 0.05 for whites and 0.027 for blacks for the initial experience effect and 0.066 for the effect of school proportion black. Given the large differential in school proportion black differential for blacks and whites, racial composition does contribute to the higher growth in the black/white achievement gap for students in the upper portion of the 3rd grade reading distribution.

5. Conclusions

This paper documents the growth in the racial achievement gap from kindergarten to 8th grade using the ECLS and Texas administrative data. The gap increases throughout the elementary school years, though the rate of growth slows with age. Decomposition into between- and within-school components shows that changes occurring between schools account for a preponderance of the increase, particularly prior to 5th grade. Although there is little systematic variation by initial achievement prior to 5th grade, following 5th grade continued increases are much larger for students with higher initial achievement and small or even negative for those at the lower end of the initial achievement distribution. Finally, although there is little or no evidence of gender differences in either the average growth in the black/white gap or changes by initial achievement, the much lower test taking rates of black boys indicates that the achievement comparisons do not capture the educational difficulties of many black boys, particularly those at the lower end of the initial achievement distribution. Moreover, the low rates of test taking for blacks and to a less extent whites in the ECLS sample among those who participate in all five waves raises the possibility that the reported growth in achievement differences understates the actual

increase during the early elementary school years.

The substantial contribution of changes between schools suggests an important role for school quality in understanding both overall black/white achievement differences and the widening gap during elementary school. We find that specific characteristics of teachers and peers previously found to have significant effects on achievement account for a sizeable portion of the growth in the achievement gap. These findings differ from those of Fryer and Levitt (2004), Murnane, Willett, Bub, and McCartney (2005) and others who do not focus on those variables for which there exist large differences by race and strong evidence that they are important determinants of achievement.[25] Although the ECLS does not lend itself to estimation of variable effects on achievement, the similarly large black/white differences in teacher experience, student turnover, and proportion black observed in the ECLS data in combination with the much more rapid growth in the achievement gap in the early grades strongly suggests that the impact of school quality differences is at least as large if not larger in the early elementary grades.

The regression results also reveal some systematic differences in variable effects by initial achievement. It appears that having a teacher with no prior experience is more costly for students with lower initial achievement, while the adverse effect on blacks of a higher proportion black appears to rise with initial achievement and explain a portion of the more rapid increase in the racial achievement gap among those with higher initial achievement.[26]

---

[25] Murnane, Willett, Bub, and McCartney (2005)find that racial composition appears to be a simple proxy for SES and that initial teacher experience has no impact, but they are unable to claim that they have credibly identified causal impacts.

[26] With respect to the teacher experience variable, it may be that teachers improve more in their teaching of lower achieving students for whom learning comes less easily, it may be that students tend to be more disruptive in schools with more lower achievers and that an important component of the return to experience is learning to manage disruption, or it may be low achievers are much more likely to attend schools that struggle to find teachers skilled in the teaching of mathematics. Our data do not permit disentangling these possible avenues for effects.

The observed patterns of growth in the black/white achievement gap and findings on the effects of specific characteristics are relevant for policies directly related to the achievement gap and those related to achievement more generally. First and most important, school quality plays an important role in the determination of achievement. Although the specific variables considered in this paper have significant effects on test scores, the overall impact of schools is almost certainly much larger. As Rivkin et al (2005) show, easily quantifiable variables do not explain the bulk of the variance in teacher and school quality (although we do not have good information about the distribution of teacher quality by race). Our analytical strategy focused entirely on identifying causal impacts, and thus a portion of the systematic influences of schools was undoubtedly ignored because we could not ensure that the relationships we saw were causal.[27]

Second, despite the strong relationship between achievement and proportion black for blacks in the Texas public schools, implications for policy remain uncertain. As Rivkin and Welch (2006) report, court decisions limit inter-district desegregation programs, and housing patterns account for the bulk of school segregation. Moreover, racial composition effects may vary by the intensity of desegregation efforts. Our sample covers a period without much in the way of new, far-reaching desegregation activity, and the relationship between achievement and racial composition might depend upon both programmatic and historical factors that determine school attendance patterns in a given district. Consequently, active initiatives designed to increase substantially black exposure to whites might produce a different relationship between achievement and racial composition.

[27] One example is the possible importance of the race match of students and teachers. Ehrenberg and Brewer (1995), Dee (2004b), and Hanushek, Kain, O'Brien, and Rivkin (2005) find that black students do better when matched with a black teacher. However, because we cannot investigate classroom linkages here, we cannot pursue this element of schools.

Table 1. U.S. Black/White Math Test Score gap (average white score minus average black score; ECLS item response theory adjusted test scores; sample weights used)

grade	fall k	spring k	1	3	5
overall	5.4	7.9	12.3	18.0	19.4
between school	4.0	6.2	10.3	15.0	16.8
within school	1.4	1.8	2.0	3.1	2.6
coeficient from school fixed eff. regression	4.5	6.0	8.8	15.8	15.6
blacks	1,165	1,276	1,283	1,249	1,249
whites	5,972	6,520	6,535	6,439	6,410

Table 2. Texas Public School Black/White Math Test Score gap (TAAS standardized test scores)

	intact cohort			entire sample		
grade	3	5	8	3	5	8
overall	0.59	0.65	0.70	0.70	0.73	0.76
between school	0.19	0.25	0.28	0.24	0.27	0.31
within school	0.39	0.40	0.42	0.46	0.46	0.44
blacks	87,941	87,941	87,941	145,282	143,524	142,073
whites	336,948	336,948	336,948	503,835	507,789	497,924

Table 3. Simulated Observed and Actual Test Score Distributions for Blacks and Whites (Pij = probability of being actual category i but observed as category j)

Initial Actual Skill Distributions

	Low	Middle	High
Blacks	0.4	0.4	0.2
Whites	0.2	0.5	0.3

Observed Test distribution

Blacks

Low	$0.4*P_{LL} + (0.4*P_{ML} + 0.2*P_{HL})$
Middle	$0.4*P_{MM} + (0.4*P_{LM} + 0.2*P_{HM})$
High	$0.2*P_{HH} + (0.4*P_{LH} + 0.4*P_{MH})$

Whites

Low	$0.2*P_{LL} + (0.5*P_{ML} + 0.3*P_{HL})$
Middle	$0.5*P_{MM} + (0.2*P_{LM} + 0.3*P_{HM})$
High	$0.3*P_{HH} + (0.2*P_{LH} + 0.5*P_{MH})$

Table 4. U.S. Black/White Math Test Score Gap by Spring Kindergarten Reading
Test Score Category (kindergarten reading score divided into 20 intervals of equal length; results reported for categories with at least 30 blacks)

	Spring Kindergarten reading test score category					
	1	2	3	4	5	6
Overall						
kindergarten	2.2	3.6	4.9	6.1	5.4	8.2
1st grade	5.7	6.4	7.8	9.7	13.7	10.4
3rd grade	14.9	12.8	15.1	13.0	15.6	9.4
5th grade	16.8	14.1	15.7	13.3	15.6	6.6
Between School						
kindergarten	2.3	3.0	3.8	5.4	5.3	7.4
1st grade	5.6	5.5	6.2	8.1	13.2	9.2
3rd grade	14.4	10.8	12.8	11.0	14.6	8.6
5th grade	16.1	11.9	13.1	11.0	14.7	6.6
blacks	208	412	324	170	65	39
whites	382	1,315	2,078	1,567	562	218

cell sizes fluctuate slightly from grade to grade

Table 5. Texas Black/White Math Test Score Gap by 3rd Grade Reading Test Score Category (students with complete observations)

	test score group, lowest to highest (-2.1 std dev to 1.2 std dev, intervals roughly .2 std dev in width)															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Overall																
3rd grade	0.51	0.49	0.43	0.41	0.37	0.41	0.35	0.32	0.30	0.30	0.27	0.24	0.22	0.21	0.18	0.16
5th grade	0.61	0.61	0.49	0.52	0.47	0.50	0.46	0.45	0.41	0.41	0.39	0.36	0.36	0.37	0.31	0.34
8th grade	0.58	0.56	0.48	0.47	0.46	0.51	0.44	0.50	0.42	0.47	0.44	0.43	0.43	0.46	0.44	0.44
Between School																
3rd grade	0.41	0.37	0.32	0.27	0.24	0.29	0.22	0.18	0.16	0.16	0.12	0.10	0.09	0.09	0.07	0.05
5th grade	0.48	0.44	0.37	0.36	0.33	0.35	0.28	0.27	0.24	0.23	0.21	0.19	0.19	0.21	0.17	0.20
8th grade	0.41	0.37	0.33	0.31	0.28	0.34	0.25	0.28	0.24	0.27	0.22	0.24	0.22	0.25	0.25	0.25
blacks	1,464	1,895	1,472	2,254	3,063	2,528	3,574	5,838	4,131	4,525	7,957	9,809	7,215	9,996	11,518	6,092
whites	1,481	2,082	1,852	2,738	4,246	3,665	6,156	10,776	8,367	10,333	21,083	31,009	29,399	51,422	82,721	61,935
overall																
3rd to 5th	0.10	0.12	0.06	0.11	0.10	0.09	0.11	0.13	0.11	0.11	0.12	0.12	0.14	0.16	0.13	0.18
5th to 8th	-0.03	-0.05	-0.01	-0.05	-0.01	0.01	-0.02	0.05	0.01	0.06	0.05	0.07	0.07	0.09	0.13	0.10
between																
3rd to 5th	0.07	0.07	0.05	0.09	0.09	0.06	0.06	0.09	0.08	0.07	0.09	0.09	0.10	0.12	0.10	0.15
5th to 8th	-0.07	-0.07	-0.04	-0.05	-0.05	-0.01	-0.03	0.01	0.00	0.04	0.01	0.05	0.03	0.04	0.08	0.05

Table 6. Estimated Effects of School and Student Characteristics on Achievement (absolute value of t statistics computed from robust std errors controlling for grouping by school in parenthesis; 1,446,458 observations)*

campus by year fixed effects	no	no	yes
campus by grade fixed effects	no	yes	yes
proportion students new to school	-0.090	-0.090	-0.069
	(5.19)	(2.83)	(2.58)
proportion students new to school*student black	-0.076	-0.013	0.003
	(2.45)	(0.53)	(0.11)
proportion students black	-0.030	-0.016	-0.025
	(1.96)	(0.30)	(0.51)
proportion students black*student black	-0.052	-0.083	-0.084
	(2.30)	(5.12)	(5.34)
Teacher Experience			
proportion 0 years	-0.045	-0.060	-0.063
	(5.45)	(5.72)	(6.40)
proportion 1 year	-0.029	-0.032	-0.040
	(3.67)	(2.97)	(3.98)
proportion 0 years*student black	-0.041	-0.038	-0.026
	(2.04)	(2.50)	(1.84)
proportion 1 year*student black	-0.003	-0.018	-0.012
	(0.12)	(1.31)	(0.99)

*All specifications include a black dummy, indicators for a transition to junior high, subsidized lunch eligibility, special education participation, and a non-structural move (all fully interacted with black), and a full set of grade by year variables.

Table 7. Black/White Differences in Teacher and Peer Characteristics for ECLS and Texas Samples

	ECLS (3rd grade)			Texas (3rd-8th grade)		
	Blacks	Whites	B-W	Blacks	Whites	B-W
proportion students new to school	0.170	0.130	0.040	0.226	0.181	0.045
teacher experience						
proportion 0 years	0.079	0.039	0.040	0.095	0.069	0.026
proportion 1 year	0.086	0.050	0.036	0.082	0.067	0.015
proportion students black				0.361	0.090	0.271
proportion of students in schools where;						
0% students are black	0.002	0.087	-0.085			
0-5% of students are black	0.034	0.544	-0.510			
5-10% of students are black	0.023	0.100	-0.077			
10-25% of students are black	0.190	0.171	0.019			
>25% of students are black	0.750	0.098	0.652			

Table 8. Contributions of Teacher Experience, Student Turnover, and Student Proportion Black to the Growth in the Black/White Achievement Gap between 3rd and 8th Grade

	Black/White difference	annual effect	cumulative effect
proportion students new to school	0.045	-0.003	-0.0085
proportion students black	0.27	-0.012	-0.0333
Teacher Experience			
proportion 0 years	0.026	-0.002	-0.005
proportion 1 year	0.015	-0.001	-0.0018
all variables combined		-0.018	-0.0486

Table 9. School by Grade and School by Year Fixed Effect Estimated Effects of School and Student Characteristics on Achievement by 3rd Grade Reading Category
(students in top four reading categories classified as high; absolute value of t statistics computed from robust std errors controlling for grouping by school in parenthesis; 1,446,458 observations)*

	coefficient	t statistic	p value for f test that effects on blacks do not vary by reading category
proportion students new to school	-0.050	(1.70)	
proportion students new to school*high	-0.024	(1.41)	
proportion students new to school*student black	-0.047	(1.76)	
proportion students new to school*student black*high	0.064	(2.17)	0.13
proportion students black	-0.059	(1.19)	
proportion students black*high	0.051	(3.32)	
proportion students black*student black	-0.019	(1.09)	
proportion students black*student black*high	-0.117	(6.05)	0.00
Teacher Experience			
proportion 0 years	-0.099	(8.14)	
proportion 0 years*high	0.050	(5.57)	
proportion 0 years*student black	0.001	(0.01)	
proportion 0 years*student black*high	-0.023	(1.24)	0.00
proportion 1 year	-0.060	(5.12)	
proportion 1 year*high	0.034	(3.88)	
proportion 1 year*student black	-0.011	(0.69)	
proportion 1 year*student black*high	0.019	(0.99)	0.14

*All specifications include black and high dummies,indicators for a transition to junior high, subsidized lunch eligibility, special education participation, and a non-structural move (all fully interacted with black), and a full set of grade by year variables.

Table A1. Grade Progression and Test Taking Status by Gender, Race, and Grade

	blacks			whites		
	3	5	8	3	5	8
All						
Has score	88.5%	85.3%	83.3%	93.5%	92.6%	90.2%
Test Missing						
in special education	9.2%	11.3%	9.2%	4.8%	4.8%	4.6%
absent during testing	1.2%	0.7%	1.1%	1.1%	0.9%	1.1%
other	0.5%	0.2%	1.5%	0.3%	0.1%	1.0%
Off grade sequence	0.6%	2.5%	4.9%	0.2%	1.6%	3.1%
boys						
Has score	85.4%	82.2%	78.3%	92.2%	91.9%	88.6%
Test Missing						
in special education	12.0%	15.0%	13.1%	6.1%	6.1%	6.3%
absent during testing	1.2%	0.7%	1.0%	1.1%	0.9%	1.0%
other	0.6%	0.2%	2.9%	0.3%	0.1%	1.4%
Off grade sequence	0.7%	1.9%	4.8%	0.3%	1.1%	2.7%
girls						
Has score	91.7%	89.8%	87.7%	94.9%	94.7%	92.8%
Test Missing						
in special education	6.2%	8.1%	7.0%	3.4%	3.6%	3.6%
absent during testing	1.2%	0.7%	1.0%	1.2%	0.9%	1.2%
other	0.4%	0.2%	1.9%	0.3%	0.1%	1.1%
Off grade sequence	0.5%	1.3%	2.5%	0.2%	0.7%	1.4%

Appendix Table a2. U.S. Black/White Math Test Score Gap by Gender (unweighted data)

	boys					girls				
grade	fall k	spring k	1	3	5	fall k	spring k	1	3	5
overall	6.3	8.6	13.8	20.0	20.4	5.8	8.1	12.1	18.9	19.9
between school	5.0	7.0	11.2	16.0	16.8	4.7	6.3	9.6	14.9	16.0
within school	1.2	1.5	2.6	4.0	3.6	1.1	1.8	2.5	4.0	4.0
blacks	587	646	648	623	631	578	630	635	626	618
whites	3,026	3,326	3,328	3,258	3,247	2,946	3,194	3,207	3,181	3,163

Appendix Table a3. Texas Public School Black/White Math Test Score gap by Gender (TAAS standardized test scores)

	boys			girls		
grade	3	5	8	3	5	8
overall	0.60	0.66	0.71	0.57	0.64	0.70
between school	0.19	0.25	0.28	0.19	0.25	0.29
within school	0.41	0.41	0.43	0.38	0.39	0.41
blacks		40,163			47,711	
whites		166,267			170,431	

Appendix Table a4. U.S. Black/White Math Test Score Gap by Gender and Spring Kindergarten Reading Test Score Category

	boys					girls				
test category	1	2	3	4	5	1	2	3	4	5
Overall										
kindergarten	2.3	4.0	5.4	7.1	3.9	1.7	2.8	4.3	5.0	6.2
1st grade	5.7	6.9	10.1	11.1	12.6	5.2	5.4	5.7	8.3	14.0
3rd grade	15.4	13.9	16.0	14.5	13.6	12.5	10.5	13.9	11.6	16.4
5th grade	16.9	14.9	15.9	14.2	16.2	15.0	12.2	15.3	12.5	15.0
Between School										
kindergarten	2.4	3.5	4.2	7.0	4.0	1.8	2.5	3.9	4.3	6.0
1st grade	5.6	6.3	8.3	10.3	12.4	5.2	4.5	5.3	7.0	13.7
3rd grade	14.7	12.7	14.5	13.0	13.0	12.3	9.2	12.3	9.9	16.4
5th grade	16.8	13.4	14.2	12.5	15.9	14.4	10.4	13.5	11.0	15.0
blacks	123	208	149	85	25	85	204	175	85	40
whites	266	756	1,000	776	251	116	559	1,078	791	311

Appendix Table a5. Texas Black/White Math Test Score Gap by Gender and 3rd Grade Reading Test Score Category (intact cohort)

	reading test score group															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Boys																
Overall																
3rd grade	0.50	0.51	0.41	0.43	0.38	0.41	0.34	0.31	0.33	0.30	0.25	0.25	0.20	0.21	0.19	0.16
5th grade	0.58	0.57	0.43	0.43	0.40	0.46	0.42	0.41	0.37	0.38	0.37	0.34	0.34	0.34	0.30	0.34
8th grade	0.61	0.61	0.49	0.43	0.40	0.49	0.42	0.45	0.40	0.43	0.43	0.42	0.42	0.44	0.43	0.44
Between																
3rd grade	0.43	0.41	0.32	0.31	0.27	0.32	0.23	0.18	0.20	0.19	0.13	0.12	0.08	0.09	0.09	0.06
5th grade	0.48	0.46	0.37	0.32	0.35	0.34	0.30	0.27	0.24	0.26	0.22	0.19	0.20	0.20	0.17	0.20
8th grade	0.49	0.47	0.35	0.32	0.29	0.35	0.27	0.29	0.24	0.27	0.23	0.24	0.23	0.25	0.25	0.25
blacks	781	960	788	1,171	1,559	1,223	1,804	2,911	1,980	2,156	3,696	4,365	3,119	4,136	4,594	2,361
whites	891	1,215	1,081	1,587	2,433	2,085	3,487	6,046	4,521	5,578	11,143	16,015	14,802	24,835	38,339	27,540
2. Girls																
Overall																
3rd grade	0.48	0.44	0.42	0.36	0.31	0.38	0.34	0.32	0.26	0.27	0.26	0.21	0.21	0.20	0.17	0.15
5th grade	0.61	0.62	0.52	0.58	0.50	0.50	0.47	0.47	0.42	0.42	0.39	0.37	0.36	0.39	0.32	0.34
8th grade	0.54	0.52	0.46	0.50	0.49	0.51	0.45	0.52	0.43	0.49	0.43	0.44	0.41	0.47	0.45	0.42
Between																
3rd grade	0.47	0.39	0.36	0.29	0.25	0.29	0.24	0.20	0.16	0.17	0.13	0.09	0.09	0.09	0.07	0.05
5th grade	0.49	0.47	0.40	0.44	0.35	0.40	0.31	0.29	0.28	0.24	0.23	0.21	0.21	0.23	0.19	0.21
8th grade	0.44	0.35	0.31	0.36	0.31	0.37	0.27	0.30	0.25	0.27	0.23	0.24	0.23	0.26	0.25	0.25
blacks	679	932	684	1,080	1,502	1,303	1,765	2,925	2,149	2,369	4,251	5,437	4,093	5,857	6,916	3,728
whites	589	867	769	1,150	1,806	1,577	2,661	4,720	3,841	4,742	9,920	14,967	14,568	26,556	44,337	34,356

Table A6. Racial Gap (black minus white) in Peer and School characteristics by 3rd Grade Reading Test Category

	reading test score group															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Peer Variables																
proportion new to school	0.049	0.039	0.032	0.030	0.038	0.036	0.036	0.035	0.033	0.035	0.036	0.036	0.038	0.038	0.041	0.042
proportion black	0.35	0.32	0.29	0.30	0.30	0.30	0.29	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.29	0.30
teacher experience																
proportion 0 years	0.047	0.029	0.027	0.025	0.025	0.023	0.021	0.026	0.021	0.021	0.023	0.023	0.022	0.025	0.024	0.028
proportion 1 year	0.013	0.019	0.013	0.009	0.008	0.012	0.012	0.012	0.012	0.011	0.012	0.015	0.013	0.014	0.015	0.015

Appendix Table a7. Estimated Effects of School and Student Characteristics on Achievement by 3rd Grade Reading Test Score Categories with at least 40,000 observations
(absolute value of t statistics computed from robust std errors controlling for grouping by school in parenthesis)

Reading Category	8	9	10	11	12	13	14	15	16
proportion students new to school	-0.17	-0.18	-0.08	-0.04	-0.10	-0.19	-0.09	-0.01	-0.29
	(1.47)	(1.42)	(0.69)	(0.52)	(1.63)	(3.18)	(2.10)	(0.28)	(1.00)
proportion students new to school*student black	-0.07	-0.23	-0.06	0.05	-0.04	0.04	0.03	-0.01	0.01
	(0.68)	(2.35)	(0.59)	(0.71)	(0.64)	(0.67)	(0.74)	(0.21)	(0.27)
proportion students black	-0.04	-0.31	0.24	-0.28	0.04	-0.08	-0.07	-0.08	-0.12
	(0.18)	(1.36)	(1.22)	(2.02)	(0.32)	(0.73)	(0.73)	(1.28)	(1.88)
proportion students black*student black	-0.02	-0.13	-0.03	-0.02	-0.04	-0.11	-0.11	-0.07	-0.11
	(0.37)	(1.83)	(0.46)	(0.47)	(1.28)	(2.98)	(3.60)	(2.88)	(3.92)
Teacher Experience									
proportion 0 years	-0.10	-0.19	-0.10	-0.07	-0.10	-0.04	-0.08	-0.06	-0.03
	(2.33)	(3.74)	(2.29)	(2.39)	(4.35)	(1.73)	(4.60)	(5.04)	(2.73)
proportion 1 year	-0.05	-0.12	-0.08	-0.01	-0.06	-0.06	-0.05	-0.03	-0.01
	(1.12)	(2.37)	(1.68)	(0.48)	(2.78)	(2.44)	(2.67)	(2.13)	(0.95)
proportion 0 years*student black	-0.03	0.08	0.01	-0.01	0.04	-0.09	-0.01	0.00	-0.05
	(0.61)	(1.28)	(0.15)	(0.34)	(1.21)	(2.42)	(0.39)	(0.00)	(1.65)
proportion 1 year*student black	0.06	0.12	0.01	0.06	-0.03	0.05	0.02	-0.01	-0.03
	(1.03)	(1.82)	(0.24)	(1.68)	(0.87)	(1.66)	(0.69)	(0.54)	(1.13)
observations	53,722	41,145	48,993	96,936	137,905	124,691	212,011	328,905	238,393

*All specifications include a black dummy,indicators for a transition to junior high, subsidized lunch eligibility, special education participation, and a non-structural move (all fully interacted with black), and a full set of grade by year variables.

Figure 1. 3rd to 5th Grade and 5th to 8th Grade Changes in the Math Achievement Gap by 3rd Grade Reading Category

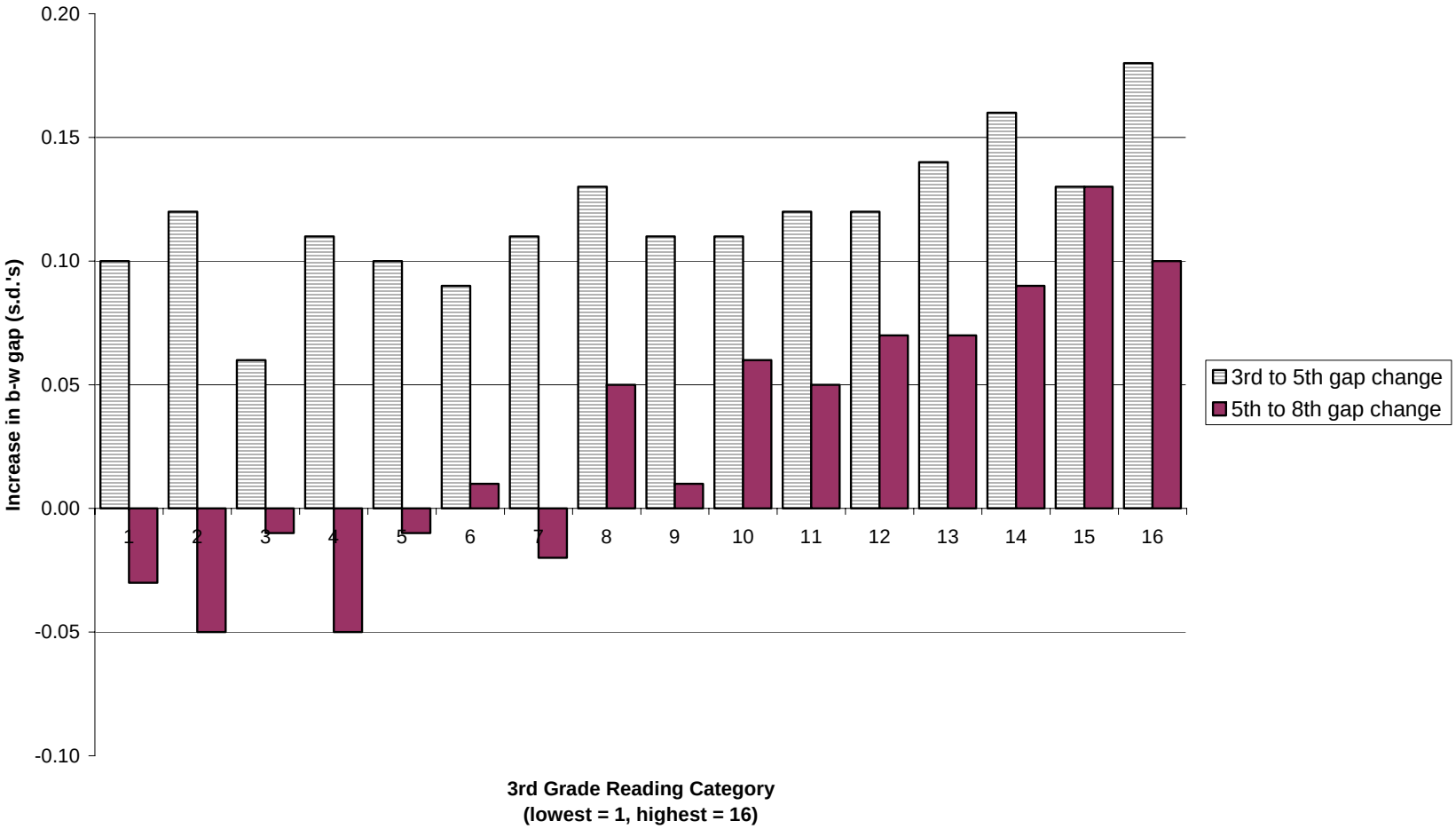
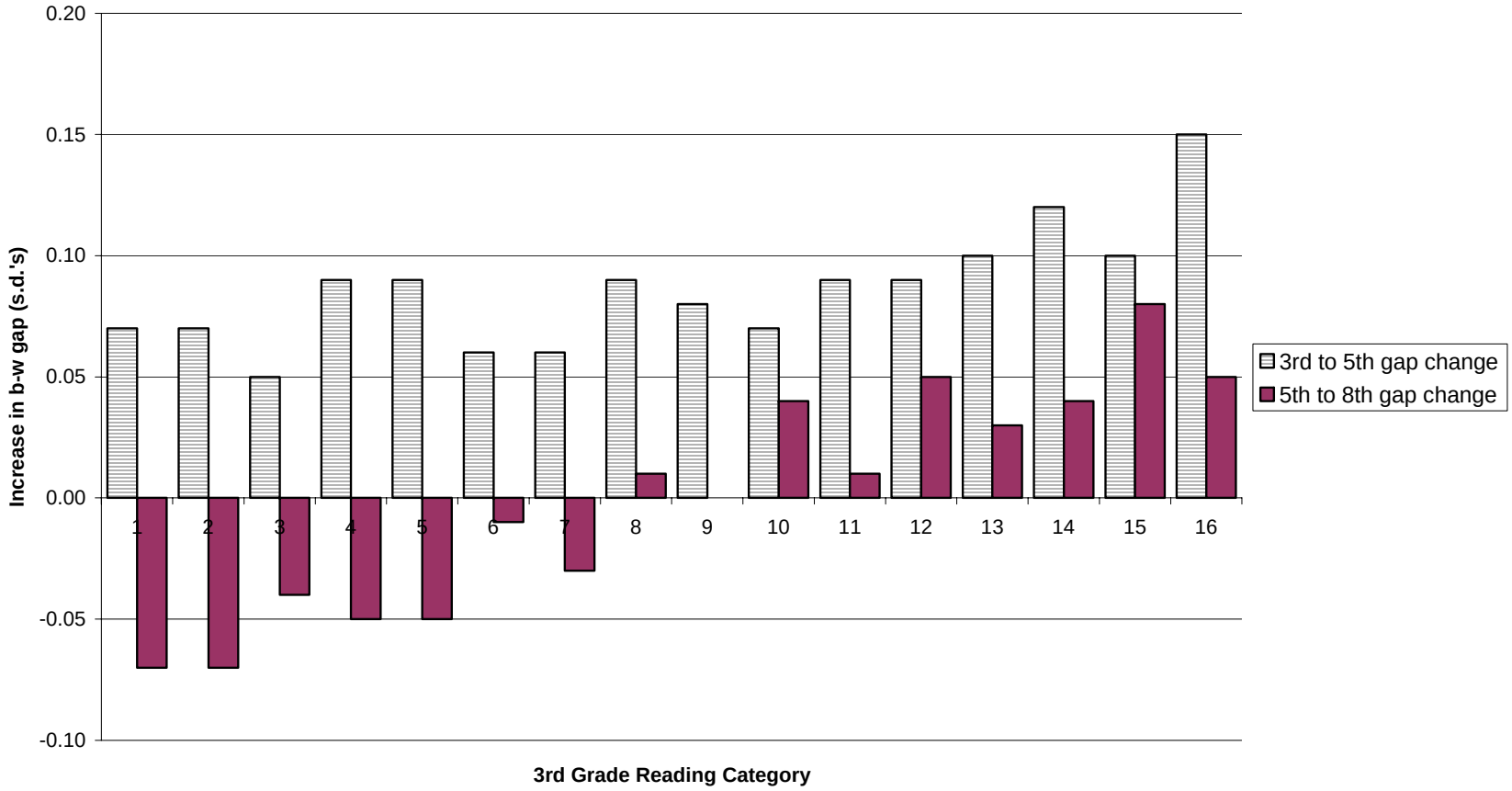


Figure 2. Between School Changes in the Math Achievement Gap by 3rd Grade Reading Category



Appendix

This appendix develops the decomposition presented in equation (1). We begin by expressing the average achievement for blacks (and whites) as equal to the weighted average of school average achievement for the respective group d:

$$\overline{A}_w - \overline{A}_b = \sum_s \frac{n_{ws}}{n_w}\overline{A}_{ws} - \sum_s \frac{n_{bs}}{n_b}\overline{A}_{bs}$$

Addition and subtraction of the overall school average achievement of all blacks and whites $\overline{A}_s$ from each term yields

$$\overline{A}_w - \overline{A}_b = \sum_s \left[\frac{n_{ws}}{n_w}\overline{A}_s + \frac{n_{ws}}{n_w}\left[\overline{A}_{ws} - \overline{A}_s\right] \right] - \sum_s \left[\frac{n_{bs}}{n_b}\overline{A}_s + \frac{n_{bs}}{n_b}\left[\overline{A}_{bs} - \overline{A}_s\right] \right]$$

Reorganizing terms produces between and within school components.

$$\overline{A}_w - \overline{A}_b = \sum_s \left[\frac{n_{ws}}{n_w}\overline{A}_s - \frac{n_{bs}}{n_b}\overline{A}_s \right] + \sum_s \left[\frac{n_{ws}}{n_w}\left[\overline{A}_{ws} - \overline{A}_s\right] - \frac{n_{bs}}{n_b}\left[\overline{A}_{bs} - \overline{A}_s\right] \right]$$

Expressing the school average achievement components $\overline{A}_s$ in the second summation term equal to the weighted average of the average achievement levels of blacks and whites in the school yields

$$\overline{A}_w - \overline{A}_b = \sum_s \left[\frac{n_{ws}}{n_w}\overline{A}_s - \frac{n_{bs}}{n_b}\overline{A}_s \right] + \sum_s \left[\frac{n_{ws}}{n_w}\left[\overline{A}_{ws} - \left(\frac{n_{ws}}{n_s}\overline{A}_{ws} + \frac{n_{bs}}{n_s}\overline{A}_{bs} \right)\right] - \frac{n_{bs}}{n_b}\left[\overline{A}_{bs} - \left(\frac{n_{ws}}{n_s}\overline{A}_{ws} + \frac{n_{bs}}{n_s}\overline{A}_{bs} \right)\right] \right]$$

Next, each term in brackets in the second summation term can be simplified by multiplying the denominators and numerators by n_s where necessary. Then the $1/n_w$ and $1/n_b$ terms can be moved outside the summation yielding

$$\overline{A}_w - \overline{A}_b = [\sum_s \frac{n_{ws}}{n_w}\overline{A}_s - \sum_s \frac{n_{bs}}{n_b}\overline{A}_s] + [\left(\frac{1}{n_w} + \frac{1}{n_b} \right)\sum_s (\overline{A}_{ws} - \overline{A}_{bs})\alpha_s(1-\alpha_s)n_s]$$

References

Angrist, Joshua D., and Kevin Lang. 2004. "Does school integration generate peer effects? Evidence from Boston's Metco Program." *American Economic Review* 94,no.5 (December):1613-1634.

Boardman, Anthony E., and Richard J. Murnane. 1979. "Using panel data to improve estimates of the determinants of educational achievement." *Sociology of Education* 52(April):113-121.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2003. "Segregation and resegregation in North Carolina's public school classrooms." *North Carolina Law Review* 81,no.4 (May):1463-1511.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York. 1966. *Equality of educational opportunity*. Washington, D.C.: U.S. Government Printing Office.

Dee, Thomas S. 2004a. "The race connection." *Education Next* 5,no.2 (Spring):53-59.

———. 2004b. "Teachers, race, and student achievement in a randomized experiment." *Review of Economics and Statistics* 86,no.1 (February):195-210.

Ehrenberg, Ronald G., and Dominic J. Brewer. 1995. "Did teachers' verbal ability and race matter in the 1960s? Coleman revisited." *Economics of Education Review* 14,no.1 (March):1-21.

Fryer, Roland G., Jr., and Steven D. Levitt. 2004. "Understanding the black-white test score gap in the first two years of school." *Review of Economics and Statistics* 86,no.2 (May):447-464.

———. 2005. "The black-white test score gap through third grade." WP 11049, National Bureau of Economic Research (January).

Greenberg, David, and John McCall. 1974. "Teacher mobility and allocation." *Journal of Human Resources* 9,no.4 (Fall):480-502.

Guryan, Jonathan. 2004. "Desegregation and black dropout rates." *American Economic Review* 94,no.4 (September):919-943.

Hanushek, Eric A. 2003. "The failure of input-based schooling policies." *Economic Journal* 113,no.485 (February):F64-F98.

Hanushek, Eric A., John F. Kain, Daniel M. O'Brien, and Steve G. Rivkin. 2005. "The market for teacher quality." Working Paper No. 11154, National Bureau of Economic Research (February).

Hanushek, Eric A., John F. Kain, and Steve G. Rivkin. 2004a. "Disruption versus Tiebout improvement: The costs and benefits of switching schools." *Journal of Public Economics* Vol 88/9-10:1721-1746.

———. 2004b. "Why public schools lose teachers." *Journal of Human Resources* 39,no.2:326-354.

———. 2006. *New evidence about Brown v. Board of Education: The Complex Effects of School Racial Composition on Achievement*. University of Texas at Dallas (revised March 2006).

Jencks, Christopher, and Meredith Phillips, eds. 1998. *The black-white test score gap*. Washington, DC: Brookings.

Lankford, Hamilton, Susanna Loeb, and James Wyckoff. 2002. "Teacher sorting and the plight of urban schools: A descriptive analysis." *Educational Evaluation and Policy Analysis* 24,no.1 (Spring):37-62.

Murnane, Richard J. 1981. "Teacher mobility revisited." *Journal of Human Resources* 16,no.1 (Winter):3-19.

Murnane, Richard J., John B. Willett, Kristen L. Bub, and Kathleen McCartney. 2005. "Understanding trends in racial/ethnic achievement gaps during elementary school." In *Brookings-Wharton papers on urban affairs, 2005*, edited by Gary Burtless and Janet Rothenberg Pack. Washington: Brookings.

National Center for Education Statistics. 2005. *NAEP 2004: Trends in Academic Progress, Three Decades of Student Performance in Reading and Mathematics*. Washington, D. C.: U.S. Department of Education.

Neal, Derek A., and William R. Johnson. 1996. "The role of pre-market factors in black-white differences." *Journal of Political Economy* 104,no.5 (October):869-895.

O'Neill, June. 1990. "The role of human capital in earnings differences between black and white men." *Journal of Economic Perspectives* 4,no.4 (Fall):25-46.

Rivkin, Steve G., and Finis Welch. 2006. "Neighborhood segregation and school integration." In *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: Elsevier.

Rivkin, Steven G. 1995. "Black/white differences in schooling and employment." *Journal of Human Resources* 30,no.4 (Fall):826-852.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, schools, and academic achievement." *Econometrica* 73,no.2 (March):417-458.

Schofield, Janet Ward. 1995. "Review of research on school desegregation's impact on elementary and secondary school students." In *Handbook of research on multicultural education*, edited by James A. Banks and Cherry A. McGee Banks. New York: Macmillan Publishing:597-616.

Todd, Petra E., and Kenneth I. Wolpin. 2003. "On the specification and estimation of the production function for cognitive achievement." *Economic Journal* 113,no.485 (February).