NBER WORKING PAPER SERIES

HARMING THE BEST:
HOW SCHOOLS AFFECT THE BLACK-WHITE ACHIEVEMENT GAP

Eric A. Hanushek
Steven G. Rivkin

Working Paper 14211
http://www.nber.org/papers/w14211

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2008

Support for this work has been provided by the Packard Humanities Institute. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Harming the Best: How Schools Affect the Black-White Achievement Gap
Eric A. Hanushek and Steven G. Rivkin
NBER Working Paper No. 14211
August 2008
JEL No. H4,I2,I28,J18

ABSTRACT

Sizeable achievement differences by race appear in early grades, but substantial uncertainty exists about the impact of school quality on the black-white achievement gap and particularly about its evolution across different parts of the achievement distribution. Texas administrative data show that the overall growth in the achievement gap between third and eighth grade is higher for students with higher initial achievement and that specific teacher and peer characteristics including teacher experience and peer racial composition explain a substantial share of the widening. The adverse effect of attending school with a high black enrollment share appears to be an important contributor to the larger growth in the achievement differential in the upper part of the test score distribution. This evidence reaffirms the major role played by peers and school quality, but also presents a policy dilemma. Teacher labor market complications, current housing patterns, legal limits in segregation efforts, and uncertainty about the overall effects of specific desegregation programs indicate that effective policy responses will almost certainly involve a set of school improvements beyond simple changes in peer racial composition and the teacher experience distribution.

Eric A. Hanushek
Hoover Institution
Stanford University
Stanford, CA 94305-6010
and NBER
hanushek@stanford.edu

Steven G. Rivkin
Amerst College
Department of Economics
P.O. Box 5000
Amherst, MA 01002-5000
and NBER
sgrivkin@amherst.edu

Harming the Best: How Schools Affect the Black-White Achievement Gap

By Eric A. Hanushek and Steven G. Rivkin

Perhaps no other social policy issue has been as important or as stubborn to deal with as racial gaps in economic outcomes. Black-white differences in academic attainment, occupation, and earnings, while showing some improvement over the past quarter century, have remained large. Much of the policy effort aimed at reducing these gaps focuses on public elementary and secondary schools. This emphasis hinges upon the widespread beliefs that school and peer characteristics disadvantage blacks relative to whites and that appropriate interventions can raise achievement and future life outcomes. This paper investigates the first of these beliefs through an examination of the changes in the black-white achievement gap as students progress through school. The findings suggest that the achievement gap increase across grades is larger for blacks with higher initial achievement, and that this is due primarily to stronger deleterious effects for initially high achieving blacks of attending schools with a high black enrollment share.

Differences by initial achievement in both the growth in the achievement gap and relationship to school racial composition are striking and carry important implications for the future education and earnings distributions. The expanding achievement racial gap as students progress through school is fueled by relatively constant gaps at the bottom of the black and the white distributions and a dramatically increasing gap at the top. Given the relationship between cognitive skills and economic outcomes, the truncation at the top of the black achievement distribution does not bode well for the future expansion of the number of blacks who complete college and graduate school and who enter high prestige occupations and positions of power.

We begin with a description of the evolution of achievement differences by initial achievement level, focusing on problems introduced by test measurement error induced regression to the mean. We follow this with an investigation of the contributions of school and peer

characteristics to these changes, paying particularly attention to the possibility that the importance of specific factors may differ by initial achievement. The findings indicate that both the growth in the achievement gap and the effects of specific variables differ significantly by position in the achievement distribution at entry to elementary school with the largest adverse impacts on blacks who enter being the best prepared.

1. Economic Motivation

Table 1 provides a stark picture of the black-white differences in academic, economic, and social outcomes that have survived the schooling policies of the last decades. Among men and women 20 to 24 years old, blacks are far less likely to complete or be in the process of completing college, far less likely to work, and far more likely to be in prison or other institutions. The rates of incarceration and non-employment for young black men paint a particularly dire picture.

Cognitive skills appear strongly correlated with black and white gaps in school attainment and in wages, and this has motivated 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 resources devoted to black students in the public schools. Nonetheless, racial disparities have been stubbornly resistant to policy, raising the possibility that schools really cannot be effective policy instruments.[3]

[1] O'Neill (1990) and Neal and Johnson (1996) provide evidence on wage differences, and Rivkin (1995) provides evidence on differences in educational attainment. A more general review of the economic impacts of cognitive skills on individual earnings is found in Hanushek and Wößmann (2008).

[2] *Brown v. Board of Education*, 347 U.S. 483 (1954).

[3] Neal (2006) documents black-white gaps in both quantity and quality of schooling and shows evidence that convergence of earlier periods slowed or stopped in the 1980s and 1990s.

Table 1. Distribution of 20 to 24 year olds by School Status, Employment Status, Years of Schooling, and Institutionalization Status in 2000 (percentages by Gender and Race)

	Institutionalized	High school dropout		High School Graduate				College Graduate		Total observations
				Attending school		Not Attending school				
		Not employed	Employed	Not employed	Employed	Not employed	Employed	Not employed	Employed	
Men										
Blacks	14.1%	10.3%	6.7%	12.7%	13.2%	15.0%	23.3%	1.2%	3.8%	10,459
Whites	2.7%	4.2%	9.5%	13.6%	22.5%	6.3%	29.0%	2.3%	10.0%	53,820
Women										
Blacks	0.9%	10.3%	5.6%	15.6%	20.0%	17.2%	21.3%	2.2%	7.0%	10,728
Whites	0.3%	6.4%	4.7%	13.9%	26.8%	9.8%	19.5%	3.1%	15.5%	50,664

Note: Row percentages add to 100 percent.

Source: Author calculations from Census 2000 Public Use Microdata Use Sample (PUMS).

2. Strands of Literature on the Racial Achievement Gap

The original Coleman Report (Coleman et al. (1966)) was mandated by the Civil Rights Act of 1964 and motivated by concerns about racial justice in the schools. It described a situation where the gap in student performance, measured in terms of grade-level equivalents, expanded over time. Its analysis, however, did not suggest that the achievement differences existing in 1965 had much to do with the schools but rather were driven by family and peers.[4] Its companion report, U.S. Commission on Civil Rights (1967), focused on the role of racial concentration in the schools as a primary factor in the existing achievement gaps.

This time period was also central to the development of several different policy thrusts that were directly related to achievement gaps. The U.S. Supreme Court ruling in *Brown* was slow to be implemented as many Southern districts resisted desegregation and as the Court repeatedly refined and sharpened its message of 1954. As a result, the most significant changes in school desegregation occurred in the late 1960s and during the 1970s (see Welch and Light (1987), Clotfelter (2004)). Although the policy innovations were dramatic, the research on implications of desegregation has been slower in coming. Much of the early work focused on the outcomes of desegregative acts (court orders, busing programs, and the like) and failed to separate the immediate reactions from the long term impacts. See, for example, Crain and Mahard (1978), Cook (1984), Armor (1995), and the assessment in Linn and Welner (2007).

A second policy initiative also expanded the federal role in education, particularly for disadvantaged groups. With the Elementary and Secondary Education Act (ESEA) of 1965, the federal government entered into the funding and operations of schools more forcefully than ever before, initiating compensatory education programs designed to bring up performance by

[4] This analysis was subjected to considerable criticism; see, for example, Bowles and Levin (1968) or Hanushek and Kain (1972).

disadvantaged students.[5] A variety of assessments about the impact of these compensatory programs for the disadvantaged raised questions about the effectiveness of ESEA, but it continued to be one of the most significant federal education programs.[6]

The implications of these programs and other factors on the racial achievement gap have not surprisingly received considerable attention. Studies of National Assessment of Education Progress (NAEP) scores document a lessening of the black-white gap during the 1980s and early 1990s and a roughly stable gap thereafter.[7] The impacts of specific family, school, peer, and community factors on this pattern have been examined, but the limited and often contradictory statistical evidence raises doubts about the efficacy of desegregation and other policies as mechanisms for addressing racial inequalities.[8]

Research on "summer fall back" supports the notion that increasing achievement gaps through the schooling years is largely not a product of the schools per se. This work, relying on summer and fall testing of achievement, suggests that learning during the school year might on average be the same for blacks and whites but that the amount of learning during the summer months heavily favors white students (Heyns (1978), Downey, von Hippel, and Broh (2004), Alexander, Entwisle, and Olson (2007)).

---

[5] Federal involvement in school accountability under the No Child Left Behind Act of 2001 is in fact the latest reauthorization of the ESEA statutes.

[6] Before each re-authorization, the federal government hired an outside evaluator to judge the effectiveness of the compensatory funding. Invariably, these studies found little success in terms of achievement (see Vinovskis (1999)). These findings dovetail with direct analyses of how school resources tend not to be related to student outcomes (e.g., Hanushek (2003)).

[7] See, for example, Congressional Budget Office (1986), Grissmer, Kirby, Berends, and Williamson (1994), the collection of research in Jencks and Phillips (1998), and Hanushek (2001).

[8] Earlier optimism about narrowing gaps (Jencks and Phillips (1998)) largely dissipated with new evidence that the black-white achievement gap stayed constant or even grew during the 1990s (National Center for Education Statistics (2005)). In terms of the specific policies that have been pursued, direct evidence on the benefits of school desegregation remains limited. 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 dropping out of high school, though data limitations and methodological concerns raise questions about the findings. Accumulated evidence does not provide strong support for the belief that higher expenditure typically leads to substantial improvements in the quality of instruction, particularly with regard to higher pay for teachers with a masters degree or substantial experience (Hanushek (2003)).

Other recent research generally provides additional support for that view. For example, Fryer and Levitt (2004, (2005) find that a substantial racial achievement gap exists at entry to school and increases with age but that the majority of the increase occurs within schools and is not explained by quantifiable school characteristics.[9] Clotfelter, Ladd, and Vigdor (2005) document a large third grade achievement gap in North Carolina that does not increase with schooling. Our past work, on the other hand, highlights meaningful achievement impacts of specific peer and teacher inputs whose distributions differ substantially by race, suggesting possible school based explanations of at least a portion of the black-white achievement differences.[10]

The concentration on the average achievement gap, however, may mask important effects that vary across the achievement distribution. For example, discussions of peer pressures on blacks – the "acting white" literature – raises the possibility that high achieving black students may face very different pressures than lower achieving blacks.[11] In addition, academic preparation relative to the median or mean student in the school likely affects the extent to which the curriculum approaches the ideal level of challenge for a student. These and other considerations suggest the possibility of differential responses across the initial achievement distribution to particular factors.

3. Texas Schools Project Data

The UTD Texas Schools Project (TSP) data set provides a unique stacked panel of school administrative data that allows us to track the universe of Texas public elementary and middle school students as they progress through school. For each cohort there are over 200,000 students in

[9] Note that 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. Neal (2006) finds little evidence of a growing gap past entry to school and discounts the role of schools in either creating or ameliorating any gaps.

[10] Hanushek, Kain, and Rivkin (2004) 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.

[11] See, for example, Fordham and Ogbu (1986), Cook and Ludwig (1997), McWhorter (2000), Ogbu (2003), Austen-Smith and Fryer (2005), Fryer (2006).

over 3,000 public schools. Unlike many data sets that sample only small numbers from each school, these data enable us to create accurate measures of peer group characteristics. We use data on four cohorts for grades three (the earliest grade tested) through eight. The most recent cohort attended eighth grade in 2002, while the earliest cohort attended eighth 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.[12]

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. Because the number of questions and average percent right varies across time and grades, test results are standardized to a mean of zero and variance equal to one. Because these tests cannot be used to measure knowledge growth with age, they provide no information on absolute racial differences. 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.

The student database is linked to detailed information on teachers including grade and

[12] Given the high rate of school switching, particularly among lower income and minority students, the possibility of following movers is an important asset of the data. In contrast, such mobility presents a serious sampling problem for survey data sets including the Early Childhood Longitudinal Survey (ECLS) that has been used in recent work on the racial achievement gap (Fryer and Levitt (2004, 2005) and Reardon (2008)). The descriptive analysis the ECLS data in Hanushek and Rivkin (2006) shows that blacks who move between grades 1 and 3 in the ECLS sample exhibit larger test score growth than stayers, a pattern contrary to that observed in most longitudinal data sets. This highlights the difficulty of generating a representative sample of a mobile population. More generally, sample selection problems in survey data almost certainly grow in magnitude with age.

subject taught, class size, years of experience, highest degree, race, gender, and student population served. Although individual student-teacher matches are not possible, students and teachers are uniquely related to a grade on each campus. Students are assigned the average class size and the distribution of teacher characteristics for teachers in regular classrooms for the appropriate grade, school, subject, and year.

We also exclude students with any missing grade-appropriate test observations, making two aspects of the data important for the subsequent comparisons. First, differential rates of special education placement directly affect the achievement comparisons because of the higher incidence among blacks. Second, we exclude all students retained in grade at any point between grades three and eight. Because the tests are not vertically linked across grades, it is not possible to locate a grade g test score in a grade g+1 distribution. Consequently, since blacks are more likely to be retained in grade, the resulting sample of blacks is further positively selected relative to the sample of whites.

The pattern of grade retention and exclusion from tests, shown in Appendix Table A1, has implications for the description and analysis of racial achievement gaps.[13] Specifically, because of the racial differential, it is necessary to interpret the descriptive patterns as understating the full extent of racial achievement differences, particularly differences in the shares of students in the lower tail of the distribution.

4. Racial Achievement Differences in Texas

We begin with descriptive information about the achievement gaps as students progress from third to eighth grade in Texas public schools. Because of our interest in distributional issues, however, it is necessary to discuss first the approach used to divide students by initial achievement

[13] As Hanushek and Rivkin (2006) show, the achievement gap of students who progress with their class understates black-white differences in academic progress at any point in time given the much higher rates of grade retention and missing tests due primarily to special education classification (particularly for boys) among blacks. For test taking and placement in special education, see Hanushek, Kain, and Rivkin (2002).

score. The impact of test measurement error evokes particular concern, because it is complicated by the substantial difference in initial achievement distributions by race.

A. Division by Initial Achievement

By dividing the third grade achievement distribution into quartiles, systematic variation in the change in the achievement gap between third and eighth grades across initial achievement categories can be identified. This is complicated, however, because test scores measure actual knowledge with error. This means two individuals with identical knowledge can be placed into different categories if one is lucky and the other unlucky in guessing or having been exposed to specific vocabulary words or mathematics problems. If such errors are uncorrelated over time, those initially placed in high achievement categories (partly on the basis of "good" errors) will tend to draw less positive errors in the subsequent year and have lower average achievement gains than those initially placed in lower categories.

Such regression to the mean also complicates black-white comparisons by achievement level because of race differences in the actual initial skill distributions. Table 2 illustrates the general problem using a stylized bivariate distribution of actual skill and measurement error that is randomly distributed and does not differ by race. The top panel reports the assumed distributions of actual skill for blacks and whites, where the distribution for blacks is more concentrated in the lower categories than the distribution for whites. The bottom panel describes the resulting distribution of observed test scores, where P_{ij} is the probability that somebody with true ability in category i is observed in category j. Importantly, these identical conditional probabilities are assumed for blacks and whites, based on the assumption that the distribution of test errors does not differ by race.

Given the assumptions regarding race differences in knowledge and measurement error, the bottom panel illustrates that a higher proportion of whites than blacks are misclassified into the lower observed skill category, while the opposite is true for the highest observed skill category. A higher percentage of the blacks in the lower achievement category are classified correctly, while a

Table 2. Simulated Observed and Actual Test Score Distributions for Blacks and Whites

(Pij = probability of being in actual category i but observed as category j)

Skill Category	Blacks	Whites
	Actual Distributions of Initial Skills	
Low	0.6	0.4
High	0.4	0.6
	Observed Test Distribution With Measurement Error	
Low	$0.6*P_{LL} + 0.4*P_{HL}$	$0.4*P_{LL} + 0.6*P_{HL}$
High	$0.6*P_{LH} + 0.4*P_{HH}$	$0.4*P_{LH} + 0.6*P_{HH}$

higher percentage of the whites have negative errors. In contrast, a higher percentage of the whites in the high achievement category are correctly classified, while a higher percentage of the blacks have positive errors. This implies that the average errors for blacks in each observed achievement category are higher than the average errors of whites. Because the expected error in the next period is zero if test errors leading to are uncorrelated over time, the expected achievement gain in the next period is higher for whites than for blacks throughout the observed initial skill distribution. Therefore a finding that the gap grows in each category would be expected even if there was no increase in the true knowledge differential.

The pattern illustrated in Table 2 invalidates the simple categorization of students on the basis of initial mathematics test scores. To overcome this problem, we use a test in a different subject to categorize students by initial 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 approach severs the link between initial category and expected difference in the error realizations for the initial and subsequent periods if the assumption of uncorrelated errors is true. However, the extent of teaching to the test, the probability of cheating (particularly by teachers), and the likelihood a student is ill (different subject tests are administered on adjacent days), are probably similar across subjects, raising the possibility of correlated errors and potentially invalidating this approach to categorizing students.[14]

The determination of the group cutoffs provides an additional issue to resolve. Because the distribution for whites has less dispersion than that for blacks and because there are far fewer blacks in the sample, we chose to divide the sample on the basis of the 25th, 50th, and 75th percentile of third grade test scores computed only over the sample of blacks. This leaves blacks split almost evenly into the four categories, while whites are concentrated in the highest quartile.

B. Growth in the achievement gap

[14] Notice that even if the errors are not correlated, this approach provides valid estimates of changes over time in the achievement gap and not of the magnitude of differences at any point in time.

Table 3 reports the evolution of the black-white achievement gap for all students combined and by initial test score quartile. Note the particularly large change between third and fourth grade in the top half of the distribution, suggesting that simply using the reading score for categorization is insufficient. Given both a priori concerns about correlated errors and this initial investigation, we drop third grade changes and focus on changes between grades 4 and 5 and between grades 8 and 9.

A clear pattern emerges showing similar increases between grades 4 and 5 across the initial achievement distribution but much larger increases between grades 5 and 8 in hither initial score categories. Between grades 4 and 8 the gap rises by only 0.02 standard deviations for students in the lowest initial achievement group, by 0.08 for students in the 2nd group, by 0.11 for students in the third group, and by 0.14 for students in the top initial achievement quartile. The larger increase for initially high achievers raises serious concerns about the existence of academic opportunities for some students striving to succeed and motivates our detailed investigation.

5. The Impact of Teacher and Peer Effects on Achievement Gaps

The role of schools in affecting the evolving racial achievement gap is controversial. Prior work suggests uncertainty about whether the racial gap expands or contracts with schooling. More importantly, since the earliest work on achievement gaps in Coleman et al. (1966), pinpointing any specific school influences has proven difficult. On the other hand, external factors have seemed clearer, such as the case of summer achievement fallback (cf. Downey, von Hippel, and Broh (2004), Alexander, Entwisle, and Olson (2007)).

We investigate 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 third grade achievement. Our primary goal here is to assess whether schools have a discernible impact on the growth of the racial achievement gap. We focus on teacher experience and school proportion black. These two school factors have been previously shown to be significant determinants of

Table 3. Texas Public School Mean Black-White Mathematics Test Score Gap for Intact Cohorts by Third Grade Reading Test Quartile[a]

	grade				observations	
	3	4	5	8	blacks	whites
Overall gap	0.59	0.62	0.65	0.70	89,563	344,833
Third Grade Reading quartile						
lowest	0.49	0.52	0.57	0.54	22,491	30,258
2nd	0.31	0.38	0.42	0.46	21,840	52,284
3rd	0.23	0.34	0.36	0.45	21,863	85,961
highest	0.20	0.34	0.38	0.48	22,676	173,582

Note: a. The third grade reading test score quartile is determined by the distribution of scores by blacks. All TAAS test scores are standardized scores with mean zero and standard deviation of one in each grade and year; all students must stay in appropriate grade and have valid mathematics test scores for grades three through eight.

achievement and to be distributed differently by race. Further, we have a strategy for credible identification of these effects. While other school factors may also be important, investigation of these specific factors is the natural extension of previous analyses.[15]

The potential import of these two factors is readily apparent in Table 4. Blacks are more likely than whites to have teachers with little or no experience and on average attend school with a much higher black enrollment share, regardless of initial achievement quartile.

Interestingly, there is little or no systematic variation in initial achievement of blacks across schools of differing proportion black, raising doubts about the importance of demographic composition as a determinant of achievement. However, as Table 5 shows, a pronounced relationship develops as students age. By eighth grade the average black math score equals -0.15 for students in schools that are less than 25 percent black, -0.28 in schools in which the black enrollment share lies between 25 and 50 percent, and -0.40 in schools that are majority black. Our question is how much of this divergence is caused by demographic composition and how much is caused by other school and family factors.

A. Empirical Model

Identifying the effects of teacher and peer characteristics on achievement is difficult primarily because the distribution of peer and teacher variables is not an accident but rather an outcome of government, teacher, and family choices. Their endogeneity impedes efforts to isolate exogenous variation in these variables.

[15] Note that by concentrating on these two factors we ignore other components of schools such as school leadership that are likely distributed more favorably for whites than blacks. We return to this below.

Table 4. Key Characteristics of Elementary Schools and Peers by Race and Initial Achievement Quartile

	Initial reading quartile			
	lowest	2nd	3rd	highest
Blacks				
Black schoolmates	38.2%	37.5%	37.4%	39.4%
Teachers with no prior experience	8.6%	8.3%	8.2%	8.3%
Teachers with one year of prior experience	7.8%	7.7%	7.7%	7.7%
Whites				
Black schoolmates	9.6%	9.2%	8.9%	8.7%
Teachers with no prior experience	6.6%	6.3%	6.0%	5.8%
Teachers with one year of prior experience	6.5%	6.2%	6.0%	5.8%

Table 5. Average Standardized Math Score by Grade, Race and School Racial Composition

	Percentage black			
	0-25	25-50	50-75	75-100
Blacks				
third grade	-0.15	-0.19	-0.24	-0.01
fifth grade	-0.17	-0.28	-0.31	-0.20
eighth grade	-0.15	-0.28	-0.40	-0.39
observations	20,784	13,664	7,014	8,407
whites				
third grade	0.41	0.39	0.30	0.36
fifth grade	0.40	0.34	0.28	0.21
eighth grade	0.42	0.39	0.25	0.21
observations	219,034	16,761	1,662	185

Our approach takes advantage of the stacked panel data from the Texas Schools Project to account for systematic factors related to choices by schools and parents that affect teacher and peer characteristics on the one hand and achievement on the other. The underlying value added models use the quasi-random variation in teacher and peer characteristics that remains following the removal of the multiple levels of fixed effects to identify the effects of specific peer and teacher characteristic.

Equation (1) highlights the key identification issues that must be addressed in the absence of random assignment. Here achievement (A_{iGsy}) for student i in grade G and school s in year y is modeled as a function of student, family, school, and peer factors:

$$(1) \qquad A_{iGsy} = \alpha_{iGy} + \beta X_{iGsy} + \delta S_{iGsy} + \lambda P_{iGsy} + e_{iGsy}$$

where P is peer composition, S is school quality (including teacher quality) in grade G, X is a vector of flows of contemporaneous family background during grade G, α is an individual-specific intercept specific to grade G in year y that captures the cumulative effects of family, neighborhood, and school experiences and characterizes the knowledge and skills that each student separately brings at entry to grade G, and e is a stochastic term capturing other unmeasured influences.

If P and S were uncorrelated with e and α,, OLS would yield unbiased estimates of the effects of peer and school characteristics. But as noted above, the complications inherent in the sorting of students, teachers, and administrators among schools – combined with existing evidence about the importance of each – strongly suggest that typically available variables contained in X will not account adequately for potentially confounding factors.

Our approach is to use the panel data techniques to control for student, family, school, and community factors that could potentially bias the estimated effects, leaving only exogenous variation in the variables of interest to identify the parameter estimates. Specifically, because we follow students over time and observe how different students perform in given schools, we can

remove systematic components of student, school, and peer factors by including a series of fixed effects in the model.

Perhaps the easiest way to see it is by considering the introduction of a school fixed effect – or separate intercept for each school – that captures time invariant differences in neighborhoods and schools, many of which are likely related to both achievement and school racial composition. These include school facilities, public services, community type, and working conditions that influence teacher supply.

But, because school quality undoubtedly varies over time and by grade for each school, we also includes interactions between school and both grade and year.[16] School-by-grade fixed effects capture any systematic differences across grades in a school that are common to all years, including elements of curriculum, testing programs, and grade retention policies. The school-by-grade fixed effect also accounts for the possibility that achievement and racial composition vary systematically with age due to potentially confounding factors, as would be the case if white exit from schools rises at the same age as achievement of blacks declines because of influences other than school racial composition. School-by-year fixed effects account for systematic year-to-year differences that are common to all grades in a school. These school-by-year fixed effects remove not only school trends in a very general way but also idiosyncratic variation over time. For example, changes in school administration and in neighborhood and local economic conditions that might 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") are all captured by these fixed effects without having to resort to available but imperfect proxies for each. 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.

[16] Note that we discuss both school and school-by-year fixed effects purely for expositional purposes. It is not possible to estimate both jointly.

It is also important to account for statewide factors that might be correlated with overall changes in racial composition and achievement. Grade, year, and year-by-grade fixed effects account for statewide trends in racial composition and achievement by grade and year and other factors including changes in test difficulty, accountability policies, textbooks and curriculum, and the like.

These fixed effects go significantly beyond studies that attempt to characterize differences in schools, neighborhoods, and policies by various direct measures of their attributes. Nonetheless, they do not account for grade specific, time varying factors that might be related to peer racial composition or teacher experience.

A clear and important example of possibly confounding effects would be any selective placement of students into classrooms that is correlated to both unobserved determinants of achievement on the one hand and classroom teacher or student characteristics on the other. Consider the possibility that parents who are stronger advocates for their children may be more likely to succeed in placing children with experienced teachers. Alternatively, principals may be more likely to place high achieving blacks in racially mixed classrooms. To deal with this potential problem, we restrict attention to variations in racial composition and teacher attributes at the grade rather than classroom level. This use of grade rather than classroom level information is closely related to the use of grade average percent black as an instrumental variable, although the IV estimator would also use within grade black-white differences in the classroom proportion black in the computation of standard errors.[17]

A second concern not addressed by aggregation to the grade level is the possibility that a spurious correlation is introduced between achievement and the variables of interest as a result of

[17] 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. Their descriptive analysis does not address implications for student performance, but, given our inclusion of school-by-year and school-by-grade fixed effects in regressions estimated separately by race, any such variation likely has a negligible effect on the estimates in this paper. Because our data do not include classroom identifiers, classroom differences cannot be used in this analysis.

parental decisions to switch schools. If the presence of more inexperienced teachers increases the likelihood that families leave the school, it becomes difficult to separate the influence of attrition from the true effect of teacher inexperience.

Bias from such selection emerges only if the selection is related systematically to within school-by-grade differences over time in the variables of interest and, importantly, to unobserved determinants of achievement. The fact that teacher assignments and other relevant aspects of school decisions are frequently not known until immediately prior to the beginning of school year argues against the presence of this potential source of bias. Moreover, transactions costs and the presence of multiple children in the majority of families would tend to limit family mobility in response to concerns about school quality for a single grade, even if relevant teacher and classroom assignments were known in a timely manner.[18] Nonetheless, we do account directly for student heterogeneity as we describe below.

The variation used to identify the parameter estimates can be illustrated by considering racial composition for a single school (In a more general case with multiple schools, the coefficients would reflect the average of these within-school relationships across the sample). With multiple years of data for one grade, we could use cohort differences in achievement and racial composition to identify the racial composition effect. However, unobserved changes over time could bias the estimates produced by this school-by-grade fixed effects model. Alternatively, with multiple grades of data for a single year, we could use grade differences in achievement and in racial composition to identify the racial composition effect. However, systematic differences by cohort or grade could bias the estimates produced by this school-by-year fixed effects model.

Fortunately, the availability of data for multiple years and grades permits the simultaneous inclusion of school-by-grade and school-by-year fixed effects. In addition, the presence of multiple schools in the sample allows for the inclusion of grade-by-year fixed effects that account for any

[18] As noted, however, ny potential reactions yielding different classroom assignments within a given school do not cause problems.

grade specific, statewide changes in policy, curriculum, or the difficulty of tests. Therefore the racial composition and teacher characteristic effects are identified by deviations from school averages for each grade and year.

In this framework, the remaining variation in racial composition and the other variables comes both from students switching schools and from persistent cohort-to-cohort differences reflecting natural demographic variations in cohort composition within schools. But, because either of these sources of variation may be systematically related to student and family determinants of achievement, we must account for student heterogeneity directly as well as the direct effects of mobility.

Mobility induced changes, although frequently ignored in research based on cohort comparisons, introduce potentially serious problems.[19] Hanushek, Kain, and Rivkin (2004) show that blacks are much more likely to change schools than whites and thus to contribute disproportionately to year-to-year changes in school racial composition. Moreover, the evidence shows that movers tend to have lower prior achievement. In order to purge these contaminating influences, we control directly for the effects of moving on school changers with a vector of mobility variables that allow for different effects by timing, number, and type of move.[20]

The key remaining issue is the appropriate method for controlling for student-specific heterogeneity, α_{iGY} in Equation (1). Many prior authors have emphasized the role of fixed ability differences, but concerns go considerably beyond this notion of heterogeneity. Equation (2)

[19] An identifying assumption in a number of studies that make use of cohort differences is that either raw cohort differences or differences remaining following the removal of school specific trends over time are not correlated with confounding factors. This approach, which builds on the intuition that students close in age in the same school have many similar experiences, has been used in a variety of circumstances (e.g., Ehrenberg and Brewer (1995), Ferguson and Ladd (1996), and more recently generalized by Hoxby (2000)).

[20] Indicator variables differentiate both among those moving during the summer, school year, or at least twice in the same year and among within district changes, district changes within geographic region, and moves across regions.

specifies α_{iGY} as a function of prior school, peer, and family variables in previous grades, and unobserved "ability" γ, a function of early childhood experiences, prenatal care, and heredity.[21]

$$(2) \qquad \alpha_{iGy} = \beta\sum_{g=1}^{G-1}\theta^{G-g}X_{igy} + \delta\sum_{g=1}^{G-1}\theta^{G-g}S_{igy} + \lambda\sum_{g=1}^{G-1}\theta^{G-g}P_{igy} + (\gamma_i + \sum_{g=1}^{G-1}\theta^{G-g}\gamma_i)$$

This formulation captures the cumulative effects of families, communities, and schools that establish the knowledge base at the start of grade G and therefore affect achievement at the end of grade G.[22] 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 $0 \le \theta \le 1$. 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. (Note that θ=0 corresponds to estimating models of the level of achievement along with an individual student fixed effect). 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 fourth grade achievement equals her impact on fifth grade achievement and achievement in all future grades. (θ=1 corresponds to estimating models specified in terms of ΔA, or the simple change in achievement across grades).

Equation (2) includes a mixture of time invariant and time varying differences that could potentially bias estimates of racial composition effects if not incorporated directly into the estimation. But, as can be readily seen by writing equations 1 and 2 for grade G-1, including the student's prior test score on the right hand side of equation 1 captures the cumulative effects of prior school, community, and family influences that might be systematically related to peer and

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

[22] This representation makes clear the interpretation of the various inputs (*X, S*, and *P*) in equation 1. These represent the flow of these inputs in each grade, while the cumulative inputs in equation 2, appropriately weighted, provide the stock of each input prior to grade G. At times the flows are measured by the level of specific inputs that do not change frequently, such as the educational attainment of parents, but the conceptual idea is that parents with different educational attainment provide differing flows of inputs to their

teacher variables without imposing an assumption about the value of θ. The coefficient on prior achievement is thus a direct estimate of θ.

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. 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.[23] On the other hand, if skills were measured by location in the distribution (as we do here with standardized scores), the complicated final term in parentheses could be replaced with γ_i, because ability induced differences in relative achievement would remain constant over time.[24]

The formulation with lagged achievement does not account explicitly for the contemporaneous effect of unobserved ability, and a key identifying assumption is that any variation in γ not correlated with the prior test score is orthogonal to the variation in the teacher and school characteristics that remains following the inclusion of the multiple levels of school fixed effects. We do not believe that this is a strong condition. To be violated, it must be the case that – conditional on prior score – schools or parents take actions to alter average teacher or peer characteristics for a specific grade in ways that are related systematically to unobserved ability.

In sum, the approach is aimed at eliminating in a very general way any systematic family, school, and neighborhood factors that might be correlated with the focal determinants of

---

child's learning. Moreover, with separation and new family relationships, these inputs can themselves vary over time.

[23] In testing terms this implies having vertically scaled scores that indicate skills and knowledge over time and not just measurement relative to a grade-specific norm for learning.

[24] Note that, more generally, this holds for all time invariant factors. Consequently, if the distributions of school quality and family and community environments were fixed through grade G, current characteristics would fully describe schooling, family, and community histories. Of course this would rule out the use of panel estimators and make it virtually impossible to identify the causal effects of specific factors. Moreover, the notion of constant school and teacher quality contradicts evidence of substantial student mobility and within-school variation over time in the quality of education.

achievement, i.e., with teacher quality and racial composition. The remaining exogenous variation through natural demographic differences among cohorts and inter-temporal movements in school circumstances thus allows for identification of how specific teacher and peer characteristics influence racial achievement gaps.

B. Results

The analysis focuses on the effects of racial composition and initial teacher experience; class size is also included in all specifications though the small differences by race rules out a sizeable role for class size in explaining growth in the achievement gap. We also considered other factors including teacher education, average teacher experience, and proportion of students who are Hispanic.[25] Consistent with prior work, however, no teacher education variable or measures of experience beyond the initial years were significant determinants of achievement, and their exclusion from the analysis had virtually no effect on the other coefficients. A similar result was found for the proportions of students who are Hispanic and Asian. Finally, the effects of all variables are allowed to differ by race and achievement quartile.

Table 6 reports estimates by race and initial achievement quartile and robust standard errors clustered by school for three specifications that progressively add school-by-year and school-by-grade fixed effects estimated separately for elementary and middle school grades.[26] In addition to the reported estimates, all specifications include indicators for subsidized lunch eligibility, participation in special education, female, a family initiated move, and a transition to

[25] Rivkin, Hanushek, and Kain (2005) finds that teacher experience is important in the first two years of a teaching career (but not thereafter) and that class size has small effects in earlier grades. These patterns are consistent with a number of other high-quality recent works including Rockoff (2004), Boyd et al. (2006), and Kane, Rockoff, and Staiger (2006). Hanushek, Kain, and Rivkin (2006) find increased concentration of black students has a particularly deleterious effect on black achievement. This finding is consistent with Guryan (2004), Angrist and Lang (2004), and Hanushek and Raymond (2005).

[26] In order to reduce computation time substantially we aggregate the data by race, initial achievement quartile, school, grade, and year and weight the regressions by the number of students in each cell. Given that the models are linear, it is not surprising that preliminary estimates of the weighted aggregate and student level models were quite similar in magnitude and significance. Estimates are based on a total of 445,741 observations in the elementary school regressions and 552,382 observations in the middle school regressions. Only black and white nonHispanic students who remain with their cohort and have nonmissing test scores for

Table 6. Estimated Effects of Racial Composition and Teacher Experience on Math Achievement by Race and Initial Test Score Quartile[a]

	elementary school			middle school		
campus-by-year fixed effects	no	no	yes	No	no	Yes
campus-by-grade fixed effects	no	yes	yes	No	yes	Yes
1. Effect of Proportion Black						
Blacks						
bottom quartile	-0.005	-0.122	-0.010	-0.317	-0.279	-0.245
	(0.065)	(0.077)	(0.095)	(0.084)	(0.132)	(0.141)
second quartile	0.056	-0.096	-0.077	-0.286	-0.257	-0.226
	(0.045)	(0.071)	(0.089)	(0.059)	(0.112)	(0.119)
third quartile	0.003	-0.147	-0.118	-0.301	-0.266	-0.234
	(0.039)	(0.069)	(0.087)	(0.057)	(0.115)	(0.118)
top quartile	-0.076	-0.240	-0.202	-0.402	-0.381	-0.348
	(0.040)	(0.069)	(0.088)	(0.067)	(0.111)	(0.114)
Whites						
bottom quartile	-0.166	-0.127	-0.106	-0.219	-0.232	-0.214
	(0.084)	(0.085)	(0.102)	(0.093)	(0.123)	(0.128)
second quartile	-0.057	-0.033	-0.002	-0.189	-0.226	-0.211
	(0.060)	(0.072)	(0.090)	(0.075)	(0.112)	(0.118)
third quartile	-0.017	0.007	0.035	-0.068	-0.111	-0.091
	(0.044)	(0.067)	(0.085)	(0.056)	(0.109)	(0.118)
top quartile	-0.011	0.038	0.069	-0.058	-0.078	-0.060
	(0.029)	(0.065)	(0.084)	(0.042)	(0.105)	(0.115)

Table 6 (cont.)

	elementary school			middle school		
campus-by-year fixed effects	no	no	yes	no	no	Yes
campus-by-grade fixed effects	no	yes	yes	no	yes	yes
2. Proportion of Teachers with 0 Yrs Experience						
Blacks						
bottom quartile	-0.155	-0.134	-0.126	-0.125	-0.057	-0.106
	(0.064)	(0.060)	(0.061)	(0.060)	(0.045)	(0.047)
second quartile	-0.137	-0.080	-0.074	-0.115	-0.077	-0.128
	(0.049)	(0.044)	(0.050)	(0.048)	(0.039)	(0.040)
third quartile	-0.222	-0.197	-0.178	-0.067	-0.039	-0.085
	(0.058)	(0.057)	(0.047)	(0.042)	(0.034)	(0.036)
top quartile	-0.191	-0.153	-0.087	-0.003	0.015	-0.033
	(0.076)	(0.077)	(0.043)	(0.044)	(0.034)	(0.037)
Whites						
bottom quartile	-0.302	-0.252	-0.268	-0.133	-0.124	-0.146
	(0.053)	(0.049)	(0.053)	(0.042)	(0.034)	(0.035)
second quartile	-0.175	-0.141	-0.153	-0.049	-0.041	-0.064
	(0.036)	(0.031)	(0.036)	(0.026)	(0.022)	(0.024)
third quartile	-0.122	-0.091	-0.107	-0.044	-0.044	-0.063
	(0.025)	(0.023)	(0.030)	(0.023)	(0.018)	(0.021)
top quartile	-0.082	-0.054	-0.069	-0.032	-0.034	-0.054
	(0.016)	(0.018)	(0.026)	(0.140)	(0.014)	(0.018)

Table 6 (cont.)

campus-by-year fixed effects	no	no	yes	no	no	yes
campus-by-grade fixed effects	no	yes	yes	no	yes	yes

3. Proportion of Teachers with 1 Yr Experience

Blacks						
bottom quartile	-0.222	-0.158	-0.110	-0.180	-0.113	-0.143
	(0.071)	(0.065)	(0.070)	(0.062)	(0.048)	(0.051)
second quartile	-0.088	-0.026	0.005	-0.062	-0.003	-0.031
	(0.050)	(0.047)	(0.050)	(0.054)	(0.039)	(0.042)
third quartile	-0.045	0.025	0.034	-0.090	-0.050	-0.074
	(0.045)	(0.045)	(0.045)	(0.048)	(0.035)	(0.036)
top quartile	-0.033	0.049	0.053	-0.073	-0.026	-0.054
	(0.040)	(0.048)	(0.042)	(0.049)	(0.031)	(0.033)
Whites						
bottom quartile	-0.093	-0.102	-0.078	0.002	0.010	-0.011
	(0.050)	(0.047)	(0.049)	(0.041)	(0.036)	(0.038)
second quartile	-0.027	-0.050	-0.034	-0.048	-0.039	-0.053
	(0.037)	(0.034)	(0.036)	(0.028)	(0.022)	(0.025)
third quartile	-0.038	-0.050	-0.035	-0.042	-0.035	-0.050
	(0.026)	(0.023)	(0.026)	(0.024)	(0.018)	(0.022)
top quartile	-0.024	-0.028	-0.016	-0.040	-0.033	-0.051
	(0.015)	(0.018)	(0.021)	(0.016)	(0.014)	(0.020)
observations		445,741			552,382	

a. Robust standard errors clustered by school in parenthesis. All specifications include a black and female indicators, 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.

middle school (in the middle school regressions), the share of students who are new to the school, and the share of teachers with two years of experience, all fully interacted by initial achievement quartile and race.

The results reported in the first panel of Table 6 reveal substantial differences in the effects of proportion black by initial achievement quartile, race, and schooling level. Consider first the two lower achievement quartiles. Estimated effects of school proportion black on blacks and whites are quite similar in both elementary and middle school. In the elementary school regressions, the hypothesis that racial composition has no effect on elementary school achievement cannot be rejected at conventional levels for either blacks or whites. In the middle school regressions, by comparison, the effects are larger and more precisely estimated; they remain quite similar in magnitude for blacks and whites, significant at the five percent level in the specification with school-by-grade fixed effects and only slightly smaller and almost significant at the five percent level in the model that also includes school-by-year fixed effects. The fact that the addition of school-by-year fixed effects causes slight reductions in coefficient magnitudes and slight increases in standard error magnitudes is consistent with possibility that the additional fixed effects exacerbate any measurement error induced attenuation bias.

In contrast, there are substantial differences across race in the estimates for students in the two higher quartiles, particularly those with the highest initial scores. The full fixed effect estimates of proportion black on achievement for blacks in the top initial achievement quartile are -0.20 (significant at the 5 percent level) in elementary school and -0.35 (significant at the one percent level) in middle school. The estimated effects for whites in the highest initial quartile, however, are small and statistically insignificant in all specifications with campus-by-grade fixed effects.

The pattern of proportion black estimates suggests peer influences that are largest for high

---

grades three through eight are included in the sample. A small number of observations are excluded because of missing information on teachers.

achieving blacks. Not only does the stability of the estimates across specifications support the notion that the estimates identify a causal effect, but it also suggests that the underlying mechanism relates to peers and not to school factors such as curriculum or school quality. Given that only grade specific variation over time is used to identify the estimates, it is unlikely that grade specific school or teacher quality that is correlated with proportion black in a particular grade and year is driving the estimates. The finding of larger effects in middle school and for initially high achieving blacks is also consistent with the beliefs that peer influences grow as students enter adolescence and that high achieving blacks come under pressure not to achieve, though the precise causal mechanism cannot be understood in the absence of more detailed information on peer interactions.

The next two panels show a strong negative relationship between achievement and the share of teachers with little or no experience. All coefficients for the share with no prior experience reported in the top panel are negative; many are significant at the five percent level for blacks, and all are significant at the five percent level for whites. For whites, a similar pattern to that observed for the peer proportion black estimates appears, as the coefficient magnitudes decline in both elementary and middle school as initial achievement rises, though in this case estimated effects tend to be larger in the younger grades. No such clear ordering appears for blacks, though the lack of precision in the estimates may conceal a true underlying pattern.

The impact of teachers with only one year of prior experience (panel 3) tends to be smaller and less precisely estimated. Regardless of initial achievement, all of the coefficients for whites in the elementary school regressions are much smaller in magnitude, and none are significant at the five percent level. For whites in the middle school, estimated effects of teachers with one year experience are quite similar to the corresponding coefficients for those with zero years of experience, with the exception of students in the bottom initial achievement category for whom the estimate is quite small and imprecise. The impact for black students in middle school tends to be similar to that of teachers with zero years of experience, but these effects are quite imprecisely estimated for elementary school.

Implications of School Effects

The combination of disparate circumstances in black and white schools and of strong and racially divergent impacts of the included school and peer variables suggests that schools account for at least a portion of the growth in the racial achievement gap. Blacks typically attend schools with a much higher black enrollment share and, as often discussed, attend schools with a larger portion of rookie teachers. To understand the impact of these on achievement, we use the estimates from the full fixed effect model reported in Table 6 to simulate the impact of setting the racial composition and teacher distribution faced by black students to the average in all of Texas schools. These simulations assume a rate of knowledge depreciation of 0.3, roughly equal to one minus the coefficient on lagged achievement score for blacks.

Table 7 shows the simulated impact on the achievement of blacks from reducing the proportion of schoolmates who are black, the proportion of teachers with no prior experience, and the proportion of teachers with one year of experience from their actual levels to the average for whites and blacks combined in the state of Texas. Other than for the lowest initial achievement group that experienced a very small increase in the gap, the simulations suggest that race differences in school proportion black and teacher experience explain between 25 and 50 percent of the growth in the achievement gap between grades four and eight. School proportion black has a much larger impact than does teacher experience for all groups, though it is important to recognize that unmeasured differences in teacher quality may be important as well.

Indeed, the impact of schools is almost certainly much larger than we show here: As Rivkin, Hanushek, and Kain (2005) indicate, easily quantifiable variables do not explain the bulk of the variance in teacher and school quality. Our analytical strategy focuses entirely on identifying causal impacts of a small number of variables, and thus a portion of the systematic influences of schools and peers were undoubtedly ignored because of impediments to the measurement and

Table 7. Simulated Effect on Average Achievement for Blacks of Reducing School Proportion Black and Proportions of Teachers With Zero or One Year of Experience to the Averages for Blacks and Whites, By Initial Achievement Quartile

	Equalizing Change in Characteristic		Annual Achievement Effect		Cumulative effect through eighth grade	% reduction in gap growth
grades	Elementary	Middle	Elementary	Middle		
1. School proportion black						
Lowest quartile	0.066	0.057	0.0007	0.0140	0.031	
low middle	0.066	0.057	0.0050	0.0130	0.031	
high middle	0.066	0.057	0.0078	0.0133	0.032	
Highest quartile	0.071	0.060	0.0144	0.0209	0.051	
2. Proportion of teachers with 0 yrs experience						
Lowest quartile	0.0046	0.0056	0.0006	0.0006	0.0015	
low middle	0.0046	0.0063	0.0003	0.0008	0.0019	
high middle	0.0051	0.0060	0.0009	0.0005	0.0014	
Highest quartile	0.0058	0.0072	0.0005	0.0002	0.0007	
3. Proportion of teachers with one yr experience						
Lowest quartile	0.0030	0.0028	0.0003	0.0004	0.0010	
low middle	0.0035	0.0030	0.0000	0.0000	0.0002	
high middle	0.0039	0.0032	-0.0001	0.0002	0.0005	
Highest quartile	0.0044	0.0032	-0.0002	0.0002	0.0003	
TOTAL						
Lowest quartile					0.033	150%
low middle					0.032	38%
high middle					0.034	27%
Highest quartile					0.052	38%

identification of variable effects.[27]

Our estimation approach relies entirely on within-school variation (across grades and years) in the teacher and peer variables. We then combined these estimates with between-school differences in the specific factors to simulate their contributions to the achievement differential. As noted earlier, we believe that other differences in teachers and schools also tend to favor whites over blacks and widen the achievement gap. One hint at the potential magnitude comes from decomposing the growth in the achievement gap into between-school and within-school components. Although the sharp divergence in school attendance patterns and accompanying uncertainty about the appropriate method for weighting observations introduces some uncertainty, several different decomposition algorithms suggest that roughly two-thirds of the growth in the gap occurred between schools for all four groups. Of course family and community differences across schools also contribute to the between-school component, so this estimate should not be interpreted as the contribution of schools per se. [28]

The Policy Dilemma

By any measure, black-white differences in schooling outcomes are a matter of enormous concern. The early progress toward racial convergence that followed *Brown v. Board of Education* and the civil rights legislation of the 1960s has slowed if not stopped over the past two decades (Neal (2006)). The implications of this slowdown for earnings inequality and the economic well-being of blacks have been magnified by the substantial increase in the return to skill experienced over the past 30 years. The differences in measured skills between blacks and whites are enormous. By age 17, the average black student is performing at around the 20th percentile of

[27] One example is the possible importance of the race match of students and teachers. Ehrenberg and Brewer (1995), Dee (2004), 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.

[28] In earlier grades, Fryer and Levitt (2006) estimate a smaller between-school share.

the white distribution.[29] This performance feeds directly into further schooling and into the labor market, contributing to substantial black-white differences in lifetime earnings and occupational prestige.

A central component of public policy to address this situation is to use schooling investments to increase the skills of minorities and disadvantaged populations and thus to ameliorate the disparate economic outcomes. To this end, two aspects of the current situation – vividly depicted in Figure 1 – are particularly telling. First, the black-white achievement gap appears to increase, not decrease, with schooling. Second, the observed gap grows most for blacks who start out at higher achievement levels.

Nonetheless, implications for policy remain uncertain. At first blush, effective policy initiatives would seem straightforward. Blacks suffer from a concentration of new teachers and from the racial concentrations in the schools they attend. Further, because black students are differentially affected by these, particularly in the case of racial composition, there appears to be room for Pareto improving policies – since according to the estimates black students can be helped at little or no cost to white students. However, Supreme Court decisions, housing patterns, and evidence regarding teacher labor markets suggest limited policy options with regard to altering student racial composition and potentially high costs to altering the distribution of inexperienced teachers.

Perhaps the most easily identified policies revolve around ensuring that black students do not draw a disproportionate share of beginning teachers. However, because a substantial portion of the existing teacher experience differential appears to result from the teacher preferences related to working conditions including location and because teachers cannot be required to teach in particular districts, urban districts may have a very difficult time raising salaries high to attract experienced teachers to high poverty, high proportion black schools. (Hanushek, Kain, and Rivkin

[29] See data on the National Assessment of Educational Progress, or NAEP (National Center for Education Statistics (2005)).

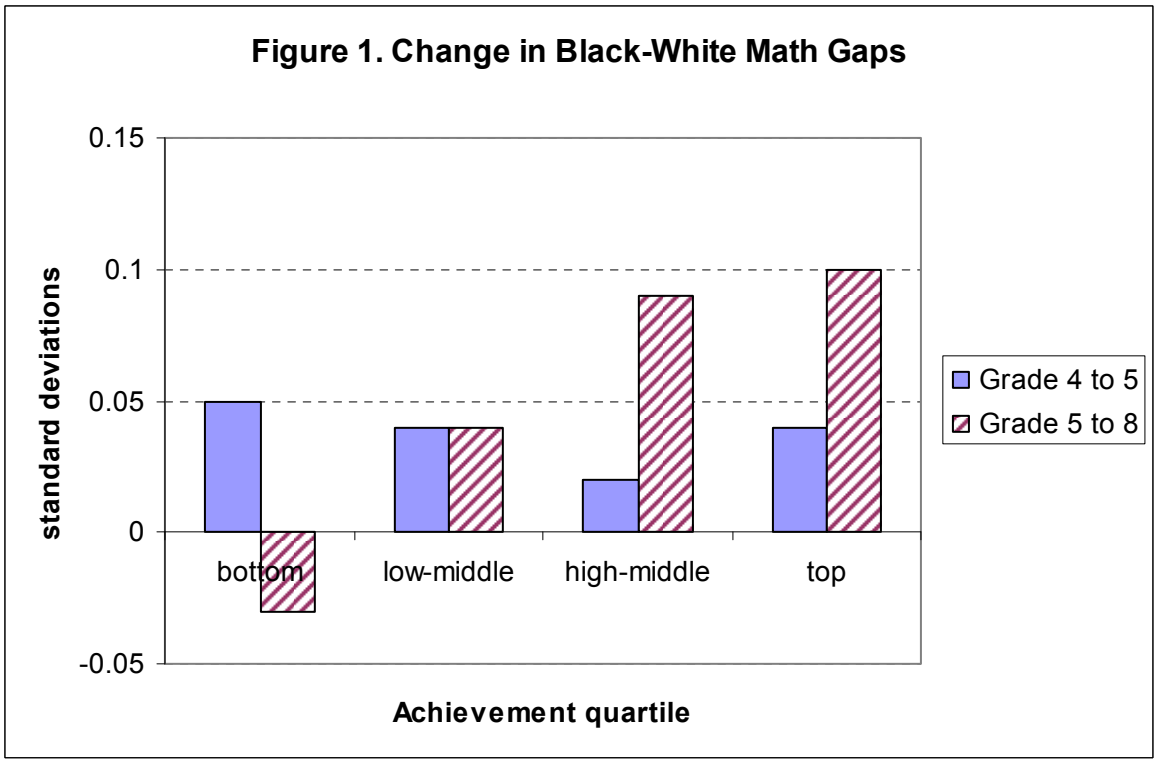
Figure 1. Change in Black-White Math Gaps
0.15
0.1
0.05
0
-0.05
standard deviations
bottom
low-middle
high-middle
top
Achievement quartile
Grade 4 to 5
Grade 5 to 8

(2004) and Boyd, Lankford, Loeb, and Wyckoff (2005)). Any changes in the process through which districts allocate teachers to schools that disadvantage experienced teachers is likely to induce exit from the district unless salaries are increased, perhaps substantially.[30]

Reducing the impact of peer composition is even more problematic. The recent U.S. Supreme Court decisions concerning schools in Seattle and Louisville severely limited if not curtailed the use of race-based policies in public schools.[31] But, more important, as Rivkin and Welch (2006) report, housing patterns account for the bulk of school segregation, and prior court decisions limit inter-district desegregation programs.[32] Moreover, our sample covers a period without much systematic 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 that we identify here.

The implication is that, while we identify specific school and peer factors that systematically affect racial achievement gaps, policy directed at just these factors is unlikely to be very successful. Instead, looking at a more comprehensive set of policies aimed at improving the quality of schools attended by blacks – such as improving teacher quality in those schools perhaps by a combination of an expansion in the pool of potential teachers through deregulation, salary increases and improvements in working conditions– will be required. Nonetheless, the magnitude of the achievement gaps is truly large, and the large gaps at entry to school point to the further need for a broader set of policies.

Finally, it is crucial to recognize that test score differences do not provide adequate information for those not in the test sample, which in this case includes students retained in grade or

[30] Hanushek, Kain, and Rivkin (2004) estimate that just equalizing the teacher outflow between central city and suburban schools in Texas would require increasing the average salary of younger teachers by over 40 percent.

[31] See Linn and Welner (2007).

those excused from test taking because of a disability or other circumstance.[33] Given the much higher rate of special education classification and grade retention for blacks than for whites and for black boys in particular, the achievement comparisons do not capture fully the gap in education progress and do not illustrate the educational difficulties of many at the lower end of the achievement distribution including those that will likely experience poor academic, social, and labor market outcomes in the future.[34]

---

[32] See *Missouri v. Jenkins,* 515 U.S. 70 (1995).

[33] See Appendix Table A1 for the quantitative magnitudes of exclusions.

[34] The low rates of test taking for blacks and to a lesser extent whites among those who participate in all five waves in the ECLS sample raise the possibility that the reported growth in achievement differences understates the actual increase during the early elementary school years.

References

Alexander, Karl L., Doris R. Entwisle, and Linda S. Olson. 2007. "Lasting consequences of the summer learning gap." *American Sociological Review* 72(April):167-180.

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.

Armor, David J. 1995. *Forced justice: School desegregation and the law*. New York: Oxford University Press.

Austen-Smith, David, and Roland G. Fryer, Jr. 2005. "An economic analysis of 'acting white'." *Quarterly Journal of Economics* 120,no.2 (May):551-583.

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.

Bowles, Samuel, and Henry M. Levin. 1968. "The determinants of scholastic achievement--an appraisal of some recent evidence." *Journal of Human Resources* 3,no.1 (Winter):3-24.

Boyd, Don, Pam Grossman, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2006. "How Changes in Entry Requirements Alter the Teacher Workforce and Affect Student Achievement." *Education Finance and Policy* 1,no.2 (Spring):176-216.

Boyd, Don, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2005. "The draw of home: How teachers' preferences for proximity disadvantage urban schools." *Journal of Policy Analysis and Management* 24,no.1 (Winter):113-132.

Clotfelter, Charles T. 2004. *After Brown: The rise and retreat of school desegregation*. Princeton, NJ: Princeton University Press.

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.

Congressional Budget Office. 1986. *Trends in Educational Achievement*. Washington, D.C.: Congressional Budget Office.

Cook, Philip J., and Jens Ludwig. 1997. "Weighing the "burden of 'acting white'": Are there race differences in attitudes toward education?" *Journal of Policy Analysis and Management* 16,no.2 (Spring):256-278.

Cook, Thomas D., ed. 1984. *School desegregation and black achievement*. Washington, DC: National Institute of Education, U.S. Department of Education.

Crain, Robert L., and Rita E. Mahard. 1978. "Desegregation and black achievement: A review of the research." *Law and Contemporary Problems* 42,no.3:17-53.

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

Downey, Douglas B., Paul T. von Hippel, and Beckett A. Broh. 2004. "Are Schools the Great Equalizer? Cognitive Inequality during the Summer Months and the School Year." *American Sociological Review* 69,no.5 (October):613-635.

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.

Ferguson, Ronald F., and Helen F. Ladd. 1996. "How and why money matters: An analysis of Alabama schools." In *Holding schools accountable: Performance-based reform in education*, edited by Helen F. Ladd. Washington, DC: Brookings:265-298.

Fordham, Signithia, and John Ogbu. 1986. "Black students' school success: Coping with the burden of 'acting white'." *The Urban Review* 58:54-84.

Fryer, Roland G., Jr. 2006. "Acting white." *Education Next* 6,no.1 (Winter):53-59.

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).

Grissmer, David W., Sheila Nataraj Kirby, Mark Berends, and Stephanie Williamson. 1994. *Student achievement and the changing American family*. Santa Monica, CA: Rand Corporation.

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

Hanushek, Eric A. 2001. "Black-white achievement differences and governmental interventions." *American Economic Review* 91,no.2 (May):24-28.

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

Hanushek, Eric A., and John F. Kain. 1972. "On the value of `equality of educational opportunity' as a guide to public policy." In *On equality of educational opportunity*, edited by Frederick Mosteller and Daniel P. Moynihan. New York: Random House:116-145.

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. 2002. "Inferring program effects for specialized populations: Does special education raise achievement for students with disabilities?" *Review of Economics and Statistics* 84,no.4 (November):584-599.

———. 2004. "Disruption versus Tiebout improvement: The costs and benefits of switching schools." *Journal of Public Economics* Vol 88/9-10:1721-1746.

———. 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).

Hanushek, Eric A., and Margaret E. Raymond. 2005. "Does school accountability lead to improved student performance?" *Journal of Policy Analysis and Management* 24,no.2 (Spring):297-327.

Hanushek, Eric A., and Steve G. Rivkin. 2006. "School Quality and the Black-White Achievement Gap." Cambridge, MA, National Bureau of Economic Research WP 12651, October).

Hanushek, Eric A., and Ludger Wößmann. 2008. "The role of cognitive skills in economic development." *Journal of Economic Literature* 46,no.3 (September).

Heyns, Barbara. 1978. *Summer learning and the effects of schooling*. New York: Academic Press.

Hoxby, Caroline Minter. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." Working Paper No. 7867, National Bureau of Economic Research (August).

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

Kane, Thomas J., Jonah E. Rockoff, and Douglas O. Staiger. 2006. "What Does Certification Tell Us About Teacher Effectiveness? Evidence from New York City." Working Paper No. 12155, National Bureau of Economic Research (April).

Linn, Robert L., and Kevin G. Welner, eds. 2007. *Race-Conscious Policies for Assigning Students to Schools: Social Science Research and the Supreme Court Cases*. Washington: National Academy of Education.

McWhorter, John H. 2000. *Losing the race: Self-sabotage in black America*. New York: The Free Press.

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. 2006. "Why has black-white skill convergence stopped?" In *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland:511-576.

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.

Ogbu, John U. 2003. *Black American students in an affluent suburb: A study of academic disengagement*. Mahwah, NJ: Lawrence Erlbaum Associates.

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.

Rivkin, Steven 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: North Holland:1019-1049.

Rockoff, Jonah E. 2004. "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data." *American Economic Review* 94,no.2 (May):247-252.

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):F3-33.

U.S. Commission on Civil Rights. 1967. *Racial isolation in the public schools*. Washington, DC: Government Printing Office.

Vinovskis, Maris A. 1999. "Do federal compensatory education programs really work? A brief historical analysis of Title I and Head Start." *American Journal of Education* 107,no.3 (May):187-209.

Welch, Finis, and Audrey Light. 1987. *New evidence on school desegregation*. Washington, D.C.: U. S. Commission on Civil Rights.

Appendix Table a1. Distribution of Texas Public School Students by Test and Grade Retention Status, by Race, Gender, and Grade

	Grade					Grade				
	4	5	6	7	8	4	5	6	7	8
	blacks					**whites**				
Girls										
has test score	88.2%	89.3%	88.6%	89.3%	87.8%	93.3%	94.3%	94.1%	94.0%	92.5%
no test score:										
special education	8.4%	9.1%	9.2%	7.5%	7.5%	4.1%	4.1%	4.1%	3.7%	4.0%
other	2.5%	1.0%	1.1%	1.6%	3.9%	2.2%	1.2%	1.2%	1.5%	3.0%
Retained in grade	0.8%	0.6%	1.1%	1.6%	0.8%	0.4%	0.4%	0.5%	0.8%	0.5%
Boys										
has test score	80.3%	81.2%	79.9%	80.8%	79.1%	90.3%	91.2%	90.6%	90.5%	88.9%
no test score:										
special education	15.9%	16.8%	16.8%	14.4%	14.3%	7.0%	7.0%	7.0%	6.4%	7.1%
other	2.7%	1.1%	1.4%	1.8%	5.2%	2.1%	1.2%	1.3%	1.6%	3.2%
Retained in grade	1.2%	0.9%	2.0%	3.1%	1.7%	0.6%	0.6%	1.1%	1.6%	1.0%

Source: Author calculations from TSP data